



# **Approaches to the diagnosis of smear-negative and sputum-scarce TB in South Africa**

**Jonathan Grant Peter, MBChB, MMed (Med), FCP (SA)**

**Thesis presented for the Degree of**

**DOCTOR OF PHILOSOPHY**

**In the Department of Medicine,**

**Faculty of Health Sciences**

**UNIVERSITY OF CAPE TOWN**

**August 2013**

**Supervisor:**

**Professor Keertan Dheda**

The copyright of this thesis vests in the author. No quotation from it or information derived from it is to be published without full acknowledgement of the source. The thesis is to be used for private study or non-commercial research purposes only.

Published by the University of Cape Town (UCT) in terms of the non-exclusive license granted to UCT by the author.

## ABSTRACT

### Background and hypotheses

In South Africa, an HIV-endemic setting, the burden of smear-negative (SN) and sputum-scarce (SS) TB is a major public health catastrophe. Diagnostic delay or failure is a key bottleneck. We hypothesised that i) assisted sputum sampling using sputum induction (SI), and ii) the use of the newer diagnostic tools (Xpert MTB/RIF and urine LAM strip), could improve diagnosis of SN or SSTB and impact clinical care. We investigated the accuracy and impact of these approaches at different levels of the health-system.

### Methods

Four cohorts were recruited. (i) In the first, overall SI diagnostic accuracy was evaluated in 696 patients with suspected SN- or SSTB referred to a hospital-based clinic. (ii) In the second, the impact of SI, performed in primary-care, on time-to-treatment was compared, using a RCT design, to an alternative simple method of healthcare worker (HCW)-provided instruction to assist sputum sampling in 517 patients. The utility of Xpert MTB/RIF was evaluated in the same context. (iii) In the third, Xpert MTB/RIF was evaluated in 496 outpatients with suspected TB. (iv) In the fourth cohort urine LAM strip test performance and incremental clinical utility was evaluated in 423 hospitalised patients with suspected TB HIV co-infection. *M. tuberculosis* liquid culture-positivity and an expanded clinical-case definition served as the primary, and where applicable secondary, reference standards.

### Results

(i) In the first cohort, SI provided a specimen  $\geq 1$ ml in 83% of patients, and culture-based diagnostic yield was 15%, with higher yields in HIV-infected versus uninfected persons (17% vs. 9%,  $p=0.02$ ).

(ii) In the second cohort, SI generated more specimens  $\geq 1$ ml (89% vs. 77%;  $p<0.001$ ) and had a higher culture-based TB diagnosis (19% vs 11%;  $p=0.02$ ) than healthcare worker (HCW)- provided instruction. However, same-day case detection using smear-microscopy (and Xpert MTB/RIF) was similar, and empiric treatment rates were

higher in instructed versus induced patients (60% vs. 38%,  $p=0.01$ ). Thus, a similar proportion in each group initiated TB treatment (25% versus 27%,  $p=0.9$ ) and there was no difference in the time-to-treatment initiation. HCW-provided instruction was less costly (US\$2.14 vs. US\$7.88) and had fewer side-effects (0% vs. 12%,  $p<0.001$ ) than SI.

(iii) In the third cohort, the sensitivity (95% confidence interval, CI) of the Xpert MTB/RIF assay was 55% (35-73) in SNTB patients using expectorated sputum specimens.

(iv) In the fourth cohort, the sensitivity (95% CI) of the urine LAM strip test was 51% (33-66) in SN- or SSTB HIV co-infected patients with  $CD4 \leq 200$  cells/ml. The ROC-curve selected grade-2, and not the manufacturer's suggested grade-1 cut-point, optimised inter-observer agreement ( $p=0.002$ ) and test specificity [grade-2: 99% (94-100) vs. grade-1: 90% (82-95),  $p=0.009$ ], thereby maximising rule-in utility (grade 2 LR+: 26.1). Urine LAM detected a similar overall proportion of TB cases to diagnosis based on clinical and CXR findings alone (45% vs. 50%,  $p=0.4$ ). However, patients detected by urine LAM, but missed by clinical and CXR screening, had higher MEWS illness severity scores ( $p=0.001$ ) and lower CD4 cell counts ( $p=0.008$ ).

## Conclusions

Sputum sampling using SI can aid SN- and SSTB diagnosis with greatest utility in HIV-infected patients. In a primary care setting high rates of empiric treatment attenuate the impact of improved SI sampling, and culture-based TB detection, on patient-important treatment outcomes. Consequently, simple HCW-provided instruction may be the preferred initial sampling strategy for adults with suspected SN- and SSTB in a resource-limited, high burden primary care setting. MTB/RIF testing on expectorated sputum specimens will substantially decrease frontline test negative TB but will not adequately address SSTB. For the latter, urine-based LAM strip testing may be a useful diagnostic adjunct to guide early treatment initiation but likely only in hospitalized TB HIV co-infected patients with advanced immunosuppression. Ongoing impact and operational studies are now required.

## ACKNOWLEDGEMENTS

The research reported in this thesis was made possible through funding support as follows: i) Discovery Health Academic Fellowship; ii) South African TB HIV training (SATBAT) research training award funded by Fogarty International Center and NIH (NIH/FIC 1U2RTW007373 and 5U2RTW007370); iii) Fogarty International Clinical Research Scholars/Fellowship that was funded through NIH grant R24TW007988; and iv) Carnegie Corporation Academic Fellowship, and v) an EDCTP TBNEAT PhD fellowship. In addition, the work received broader infrastructural support through my supervisor, Professor Keertan Dheda's grant funding. Professor Dheda is supported by the EDCTP (TB NEAT and TESA grants), SA DST and the NRF (SARChI).

I thank my supervisor, Professor Keertan Dheda, for his continued motivation, interesting and helpful academic discussion, wise advice, and detailed review and critique of manuscripts as they were developed towards publication. I would also like to thank Dr Grant Theron, a post-doctoral fellow at the UCT Lung Infection and Immunity unit, whom I worked closely with on many aspects of this project, and whose academic input and critique of my manuscripts have enhanced my scientific rigour and attention to detail.

I thank all the patients who participated in these studies and the staff at the facilities where the work was conducted (Gugulethu, Langa, and Chapel Street primary care clinics; Groote Schuur, Victoria, Somerset, and GF Jooste Hospitals). I thank all the dedicated research staff of the UCT Lung Infection and Immunity unit, particularly the many study research nurses for their dedication and care for patients, the data team for their rigorous quality control, and the laboratory staff for training me in laboratory techniques and processing and performing diagnostic tests. I thank my colleagues who participated in the clinical management and review of study patients (Johnson Thomas, Mellissa Pascoe, Asheen Haripersad, Sarah Kraus, and Lynelle Mottay). I also wish to thank the administrative staff of the UCT Lung Infection and Immunity unit for help with the many hours of paperwork required to conduct research studies.

Finally, I thank my family for their ongoing love, support and encouragement during this PhD process and for nurturing and developing my academic interests and enthusiasm. And to my wife, Joanne, I give thanks for not only supporting and encouraging me to go the distance amidst many competing demands, but also for editorial assistance and loving critique.

## ABBREVIATIONS

AFB	Acid-fast bacilli
Ag	Antigen
AIDS	Acquired immunodeficiency syndrome
ARV	Antiretrovirals
BAL	Bronchoalveolar lavage
BCG	Bacille Calmette Guérin
CFU	Colony forming unit
CI	Confidence interval
CSF	Cerebro-spinal fluid
C <sub>T</sub>	Cycle threshold
DNA	Deoxyribonucleic acid
DR	Drug-resistant
DST	Drug susceptibility testing
CNS	Central nervous system
CXR	Chest x-ray
ELISA	Enzyme-linked immunosorbent assay
EPTB	Extrapulmonary tuberculosis
FNAB	Fine needle aspirate biopsy
GIT	Gastro-intestinal tract
HAART	High active anti-retroviral therapy
HBC	High burden countries
HIV	Human immunodeficiency virus
IGRA	Interferon-gamma release assay
IPT	isoniazid preventive therapy
LAM	Lipoarabinomannan
LAMP	Loop mediated isothermal amplification
LED	Light emitting diode
LOD	Limit of detection
LR+	Positive likelihood ratio
LR-	Negative likelihood ratio
LTBI	Latent tuberculosis infection
MDR TB	Multidrug-resistant tuberculosis
MEWS	Modified early warning score

MODS	Microscopic observed drug susceptibility
MGIT	Microscopic growth-in-tube
MTB/RIF	Genexpert MTB/RIF assay
<i>M. tb</i>	<i>Mycobacterium tuberculosis</i>
NAATs	Nucleic acid amplification tests
NPV	Negative predictive value
NTM	Non-tuberculous mycobacteria
OD	Optical density
PCR	Polymerase chain reaction
PPD	Purified protein derivative
PPV	Positive predictive value
POC	Point-of-care
PTB	Pulmonary tuberculosis
RIF	Rifampicin
ROC	Receiver operating characteristic
rpoB	RNA polymerase $\beta$
SPC	Sample processing control
SOP	Standard operating procedures
SNTB	Sputum negative TB
SSTB	Sputum scarce TB (unable to spontaneously expectorate sputum)
TST	Tuberculin skin test
UCT	University of Cape Town
USA	United States of America
WHO	World Health Organisation
XDR TB	Extremely drug resistant tuberculosis
ZN	Ziehl-Neelsen

## Table of Contents

<b>ABSTRACT .....</b>	<b>2</b>
<b>ACKNOWLEDGEMENTS .....</b>	<b>4</b>
<b>Chapter 1. ....</b>	<b>8</b>
Introduction	
<b>Chapter 2. ....</b>	<b>25</b>
Background and Literature review	
<b>Chapter 3. ....</b>	<b>86</b>
Sputum induction to aid the diagnosis of smear-negative or sputum-scarce TB in adults from a HIV-endemic setting	
<b>Chapter 4. ....</b>	<b>115</b>
Comparison of two methods for acquisition of sputum samples for diagnosis of suspected tuberculosis in smear-negative or sputum-scarce people: a randomised controlled trial	
<b>Chapter 5. ....</b>	<b>124</b>
Evaluation of the Xpert <sup>®</sup> MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting	
<b>Chapter 6. ....</b>	<b>135</b>
Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients	
<b>Chapter 7. ....</b>	<b>148</b>
The diagnostic accuracy of urine-based Xpert MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce*	
<b>Chapter 8. ....</b>	<b>158</b>
The relationship between urine LAM concentration and compartment-specific TB bacillary load in hospitalised HIV-infected patients with advanced immunosuppression	
<b>Chapter 9. ....</b>	<b>163</b>
Can point-of-care urine LAM strip testing for tuberculosis add value to clinical decision making in hospitalised HIV-infected persons?	
<b>Chapter 10. ....</b>	<b>173</b>
Conclusions	
<b>Appendices .....</b>	<b>192</b>
1. PhD declaration .....	<b>192</b>
2. Chapter 4 supplementary material.....	<b>201</b>
3. Chapter 5 supplementary material.....	<b>211</b>
4. Chapter 6 supplementary material.....	<b>216</b>
5. Chapter 9 supplementary material.....	<b>232</b>

## **Chapter 1.**

### **Introduction**

University of Cape Town

## Context

The Herculean battle against tuberculosis (TB) has been underway since “modern” humans expanded out of Africa (3). However, with the colliding epidemic of HIV, the epicentre of the war against TB has shifted. In 2010, there were 8.8 million new cases of TB and ~1.5 million deaths (4). South Africa (SA) reported the third highest incidence of TB globally, with over 60% of new cases co-infected with HIV (4). SA is the country worst affected by the co-epidemics of HIV and TB – with only 0.7% of the world’s population SA has approximately one-quarter of all cases of TB HIV co-infection (5). TB remains the leading cause of death in persons with HIV infection (6) (Figure 1).

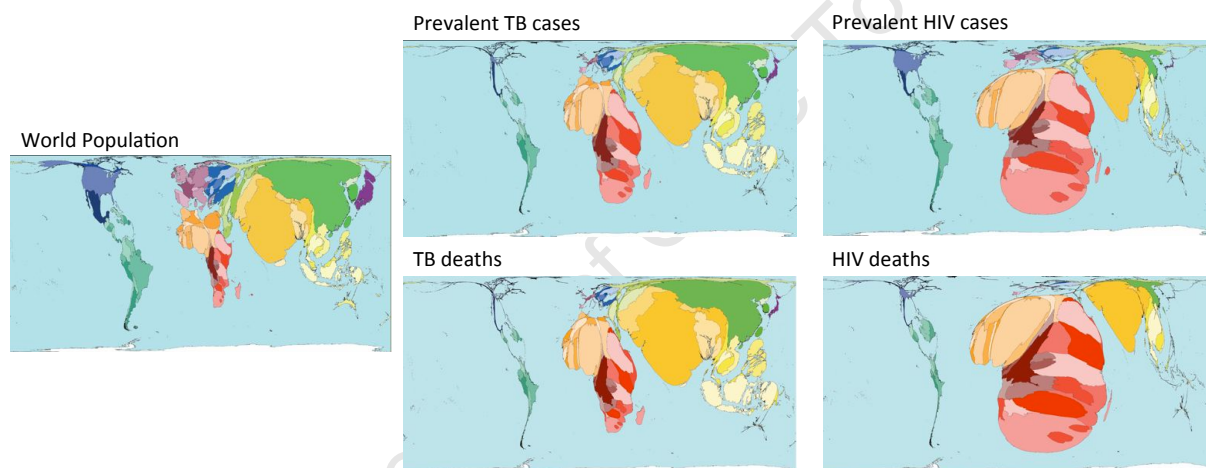


Figure 1. Re-sized world maps to highlight the co-localisation and relative global burden of TB and HIV in Sub-Saharan Africa (provided by Worldmapper with 2004 WHO statistics).

Colours represent geographical areas according to the key below:

i	Central Africa
ii	Southeastern Africa
iii	Northern Africa
iv	Southern Asia
v	Asia Pacific
vi	Middle East
vii	Eastern Asia
viii	South America
ix	Eastern Europe
x	North America
xi	Western Europe
xii	Japan

A key bottleneck of global TB control efforts is delayed or wrong diagnosis. Modeling studies suggest that the availability of more accurate diagnostic tools, especially easily accessible point-of-care tests, could decrease TB prevalence and annual TB-related mortality by as much as 20-35% (7, 8). In most of the 22 high TB burden countries, sputum smear-microscopy remains the frontline – and sometimes the only – TB diagnostic tool (9). However, this century old tool is sub-optimal, especially in high HIV prevalent settings, where sensitivities can be as low as 20% in the context of advanced immunosuppression and disseminated TB with paucibacillary pulmonary disease (10). Furthermore, up to a third of adult TB HIV co-infected cases are unable to spontaneously produce sputum for diagnostic testing (11, 12). Consequently, in HIV-endemic settings the majority of incident TB cases are smear-negative, and in SA in 2010, over 60% of notified TB cases were registered as either smear-negative or smear unknown (4). The Global Laboratory Initiative, the research community, and many national TB programmes are making ongoing efforts to address this burden, with improvements to smear microscopy and increased availability and simplification of culture-based TB diagnosis as examples (9). However, smear-microscopy remains sub-optimal, while culture-based diagnosis remains both unavailable and unaffordable in many HIV-endemic countries. Furthermore, the ability of culture-based diagnosis to impact patient-important outcomes is uncertain, given the long time delay to results (up to eight weeks). Thus, the need to develop novel approaches to SN- and SSTB, using field-friendly rapid diagnostic tools and strategies, continues to be an important public health priority in HIV-endemic high burden settings.

In the pre-HIV era, SNTB was most common in children and the elderly, and was associated with minimal non-cavitary disease (13). Both the infectivity and mortality associated with SNTB was considerably lower than smear-positive disease, and consequently it was considered to be a relatively benign form of TB (13). However, the HIV-epidemic has altered the clinical entity of SNTB (13). Although the infectivity of

SNTB appears to be unaffected by HIV status, HIV-positive patients diagnosed with SNTB have an approximately fourfold higher mortality rate than HIV-negative smear-positive TB cases (13, 14). The pauci-bacillary pulmonary disease found in SNTB patients, unlike the contained, benign disease found in HIV-negative patients, usually indicates disseminated and/or extrapulmonary TB disease in HIV-positive patients (13). Thus, rapid and more effective diagnosis of SNTB allowing for the early initiation of TB treatment could not only decrease transmission and TB incidence, but also reduce TB-related mortality. A recent hospital-based study of HIV-infected patients in Kwa-Zulu Natal, South Africa (15), which showed a reduction in 8-week mortality with the early initiation of anti-TB treatment based on the WHO smear-negative clinical algorithm, supports this hypothesis (16).

The studies reported in this thesis aimed to evaluate three approaches to improve SN- and SSTB diagnosis in a high HIV prevalence South African setting. The work has evolved beyond a simple evaluation of the diagnostic accuracy of these novel strategies and tools to investigate: i) the incremental utility of these approaches within the clinical context of high rates of routine empiric treatment based on clinical and CXR suspicion alone; and ii) the impact on patient-important outcomes. Finally, the utility of these different diagnostic approaches are considered at specific levels of the health system i.e. primary-care or hospital-based.

## **Research questions:**

The key research questions investigated in this thesis focused on addressing important research gaps in the diagnosis of adults with suspected SN- and SSTB in a high HIV TB prevalence setting. They included:

### *Improved sputum sample acquisition*

1. What is the routine performance of sputum induction (SI) in a hospital-based respiratory clinic and how does performance differ in HIV-infected versus uninfected, or out- versus in-patients?
2. What is the impact on patient important outcomes of using SI compared with simple healthcare-worker provided instruction to improve sputum sample acquisition in adults with SN- and SSTB in a primary care clinic setting?

### *Novel diagnostics*

#### *Xpert MTB/RIF assay*

3. What is the diagnostic accuracy of MTB/RIF in SNTB at the primary care clinic level and can MTB/RIF-generated Cycle threshold ( $C_T$ )-values be used as a rapid novel marker of sputum bacillary loads?

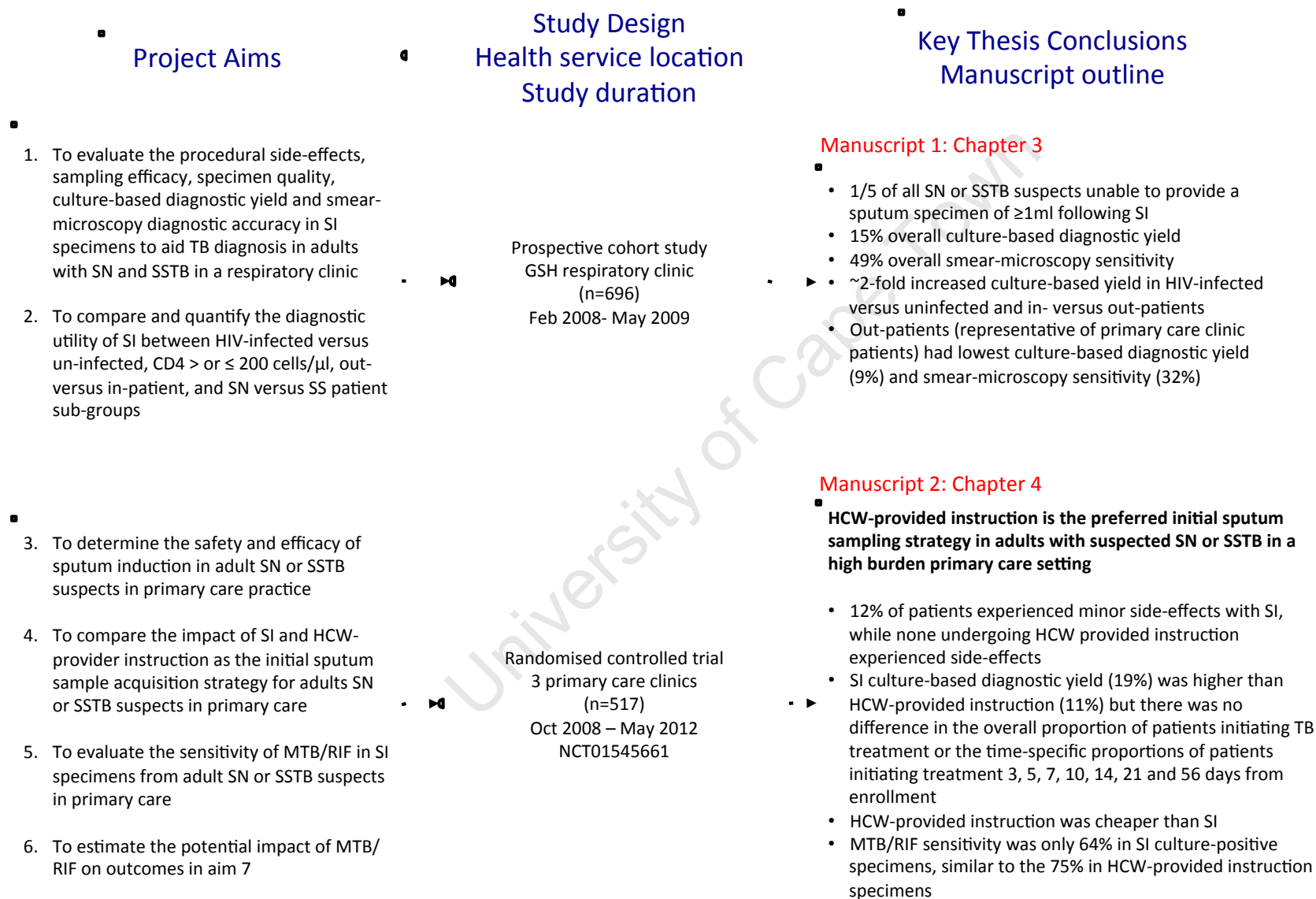
#### *Urine-based LAM strip testing*

4. What is the diagnostic accuracy of the point-of-care urine lipoarabinomannan (LAM) strip test, used alone or in combination with sputum smear-microscopy, in hospitalised HIV-infected patients with advancing immunosuppression?
5. What are the predictors of urine LAM positivity and could urine LAM concentration (or strip grade) be used as a rapid novel marker of total bacillary loads?
6. Does urine LAM strip testing offer any incremental clinical utility over routine early empiric treatment practices based on clinical and CXR parameters in hospitalised HIV-infected patients with advancing immunosuppression?

## **Overall project description**

This project arose out of discussions with my supervisor, Professor Keertan Dheda from the Division of Pulmonology, during my training as a specialist physician. As a trainee physician I was faced with the daily challenge of diagnosing TB in smear-negative or sputum-scarce HIV co-infected patients with advanced immunosuppression, and managing the devastating consequences of delayed or wrong diagnosis. These discussions lead to the development of research protocols to tackle the diagnosis of SN- and SSTB using three approaches suitable for high TB HIV burden resource-limited settings. With funding support from the Discovery Foundation and the preliminary results of the SI prospective cohort study, we developed the hypothesis for a randomised controlled trial of SI in primary care. Further funding support was secured from the NIH Fogarty International Centre and the Carnegie Corporation scheme. An overview of the PhD thesis is provided in Figure 2 below, highlighting the project aims, chronology and methodology used to address these aims, together with the key study conclusions.

Figure 2. Detailed project aims, chronology of studies, methodology, and key thesis conclusions



- 7. Evaluate diagnostic accuracy of MTB/RIF in primary care clinic TB suspects able to self-expectorate sputum, to determine potential benefits in SNTB
- 8. Investigate the correlation of MTB/RIF-generated CT-values with traditional markers of bacillary burden

Prospective cohort study  
3 primary care clinics  
(n=496)  
Feb 2007- April 2010

### Manuscript 3: Chapter 5

- MTB/RIF overall sensitivity of 55% in SNTB
- Negative predictive values of a single MTB/RIF lower in HIV-infected versus un-infected patients
- MTB/RIF-generated CT-values show good correlation with liquid culture time-to-positivity and smear grade

- 9. To evaluate the diagnostic accuracy of the point-of-care urine LAM strip test, alone or in combination with sputum smear microscopy in hospitalised patients with suspected TB HIV co-infection and advanced immunosuppression
- 10. To explore the association between mycobacteriuria and LAM-positivity in HIV-infected hospitalised patients
- 11. To investigate the correlation of urine LAM concentration with both traditional markers of bacillary burden and MTB/RIF-generated Ct-Values
- 12. To evaluate the incremental diagnostic utility of urine LAM strip testing over routine empiric treatment practice and a set of clinical predictors

Prospective cohort study  
4 district-level hospitals  
(n=423)  
July 2009- Dec 2010

### Manuscripts 4,5,6,& 7: Chapter 6,7,8,&9

- The ROC-curve selected grade-2 cut-point and not the manufacturer's suggested grade-1 cut-point offered the best specificity and inter-reader agreement
- Overall LAM strip test (grade-2) sensitivity was 50%, improving to 65% in CD4≤100 cells/μl
- LAM strip test sensitivity in HIV-infected SNTB was 39% and combined smear + LAM strip sensitivity (71%) was superior to either test alone
- Urine MTB/RIF-positivity, abnormal protein/creatinine ratio and decreasing GFR were the strongest predictors of increased urine LAM concentration
- Urine LAM concentration does not correlate with compartment-specific traditional markers of bacillary burden
- Urine LAM strip and routine early empiric treatment detected an equivalent overall proportion of HIV TB co-infected patients, but
- Urine LAM strip testing detected a group with more advanced immunosuppression and greater illness severity missed by early empiric treatment

## **Setting**

These studies were conducted in Cape Town in the Western Cape province of South Africa. The MTB/RIF prospective cohort study and the randomised controlled trial of SI enrolled patients from the Gugulethu, Langa and Chapel Street primary care clinics. These primary care clinics serve a total population of ~100 000 people, with TB incidence rates of 900-1500/100 000. In contrast, the prospective cohort study to evaluate urine LAM strip testing enrolled patients from Groote Schuur, Victoria, New Somerset and GF Jooste Hospitals. These three district-level and one tertiary-level hospital serve the greater Cape Town population of approximately two million people with an average TB incidence rate of ~900/100 000. The feasibility, performance, and impact of novel diagnostic tests varies widely depending on the setting of use. Thus, the settings of these studies were carefully considered and selected to both target the area of greatest need (e.g. the study of SI for sampling in primary care) and to optimise test performance and hence potential impact (e.g. using urine LAM strip testing in hospitalised HIV-infected patients with advanced immunosuppression).

The performance of the MTB/RIF assay, urine LAM ELISA and strip testing occurred in the research laboratories of the University of Cape Town, Lung Infection and Immunity unit in the H-floor of the Old Main Building, Groote Schuur Hospital. Smear microscopy and TB culture testing, used for all treatment decision-making, were performed by the accredited National Health Laboratory Services at Groote Schuur Hospital and Greenpoint.

## **Chronology of studies**

The first prospective cohort study began recruitment in February 2008 to evaluate the routine performance of SI at the Groote Schuur Hospital respiratory clinic. After reviewing the preliminary results of this study, we began developing the hypothesis for a randomised controlled trial to evaluate the impact of SI for adults with suspected SN- and SSTB (the SINET study) and a small pilot project was started in October 2008. Meanwhile, the prospective cohort study of SI continued to enroll patients until May 2009 in order to attain large patient numbers (n=696) to address the specific aims of comparative performance in key patient sub-groups (see Figure 2). During the development of the SINET study protocol, I became aware of the novel point-of-care urine LAM strip test. Given its obvious applicability to high HIV prevalent resource-limited settings, I decided that this would be an ideal alternative non-sputum-based diagnostic approach to evaluate for adults with SN- and SSTB, particularly targeting the vulnerable group of hospitalised patients with HIV co-infection and the highest TB-related mortality.

In 2009 I registered for full-time PhD study, and after securing funding for these research projects, we began enrolling patients in two studies. The SINET study, after the small pilot project (n=24) conducted from October 2008 to July 2009, enrolled patients from August 2009 and completed follow-up in May 2012 with a total of 517 recruited patients. The urine LAM study enrolled patients from July 2009 and completed follow-up in December 2010 with a total of 423 recruited patients. In addition, after the landmark *NEJM* publication by Boehme et al. on the MTB/RIF assay in September 2010 (2) and its likely impact on the diagnosis of smear-negative TB, we decided it essential to the relevance of my PhD work to i) evaluate MTB/RIF performance in SNTB patients recruited from the same three primary care clinics (enrolling patients in the SINET study) using stored samples from a prospective cohort of 496 patients recruited between February 2007 and

April 2009, and ii) perform MTB/RIF on stored induced sputum samples in the SINET study to investigate its potential impact on study findings.

## **Outline of thesis**

The outline of key conclusions for each manuscript and chapter of this PhD thesis is highlighted in **Figure 2**. In **Chapter 2** (Background and literature review) I present the relevant background highlighting the current challenges in TB diagnosis, the diagnostic priorities in high versus low burden settings, and the current state-of-the-art diagnosis for active TB with the focus on SN- and SSTB patients. In addition, further detailed literature is presented on sputum sample acquisition methods and studies of SI for adult SN- and SSTB patients in high HIV prevalence settings, as well as urine-based TB diagnostics and in particular urinary lipoarabinomannan (LAM).

**In Chapter 3** I present the prospective cohort study of SI to aid TB diagnosis in adults with suspected SN- or SSTB referred to Groote Schuur Hospital Respiratory clinic induction facilities. In this study, the largest programmatic cohort of patients collected to date, we evaluated the overall procedural side-effects, sampling efficacy, specimen quality, culture-based diagnostic yield and smear-microscopy diagnostic accuracy following SI. In addition, given the large numbers, we were able to compare and quantify the differences in the diagnostic performance of induced sputum in important patient sub-groups (e.g. HIV-infected versus uninfected, out-versus inpatients).

**In Chapter 4**, I report the findings of our pragmatic randomised controlled trial (RCT) to determine the preferred initial sputum sampling strategy for adults with suspected SN- or SSTB in primary care practice. This study was undertaken given

the paucity of any data, and particularly patient-important impact data, on sputum sampling strategies in primary care settings where early diagnosis may have the greatest impact. Our data presented in **Chapter 3** suggested significantly reduced performance for SI amongst outpatients, while other studies demonstrated the benefits of simple healthcare-worker provided instruction for self-expectorating primary care patients (18). Thus, using a pragmatic RCT study design we directly compared the two sampling strategies and, in one of the few diagnostic studies of this kind, we present outcomes beyond simple diagnostic accuracy measures, including the proportion of patients initiating treatment and the time-to-treatment. In addition, using stored samples, the potential impact of using MTB/RIF sputum testing is presented.

In **Chapter 5**, we present our local evaluation of the novel MTB/RIF assay in a previously collected prospective cohort of primary care clinic patients is presented. This study investigated the diagnostic accuracy of this novel tool in smear-positive and SNTB cases. In addition, we explored a number of other important research gaps pertaining to MTB/RIF use including: i) the effects of altering sputum sample volume on assay performance; ii) the relationship between MTB/RIF-generated  $C_T$ -values and traditional markers of bacterial load (measured using smear grade and culture time-to-positivity); iii) the additional yield of MTB/RIF over culture and the significance of MTB/RIF-positive, culture-negative samples; and iv) the combined diagnostic performance of smear microscopy and MTB/RIF. This manuscript was the first published independent academic evaluation of MTB/RIF in South Africa, and has already been cited more than 70 times.

In **Chapter 6** I report the findings of a prospective cohort study of hospitalised patients with suspected TB HIV co-infection and advanced immunosuppression.

The aim of this study was to conduct the first evaluation of the novel point-of-care urine LAM strip test in this patient setting. Consequently, we compared the performance of the test (including diagnostic accuracy measures and inter-reader agreement) using the manufacturer's suggested cut-point and a ROC-curve selected cut-point optimising rule-in test utility. In addition, the performance characteristics of the test in different CD4 cell count strata and used together with sputum smear microscopy are presented.

In **Chapter 7** I present data on the performance of urine-based MTB/RIF in the prospective patient cohort described in **Chapter 6**. The aim of this study was to evaluate the diagnostic utility of urine-based MTB/RIF in sputum-scarce patients with HIV-infection. Furthermore, we explored the relationship between LAM-positivity and mycobacteriuria (as determined by urine MTB/RIF-positivity) and examined other determinates of mycobacteriuria.

In **Chapters 8 and 9**, I present further analyses from the same patient cohort to more comprehensively explore and address the incremental clinical utility of urinary LAM testing. In **Chapter 8** we continue the investigation of urine LAM-positivity from an alternative perspective to that presented in **Chapter 7**, exploring the predictors of increasing urine LAM concentration (or strip grade) and the relationship of urine LAM concentration to both the traditional markers of bacillary burden mentioned previously and urine-based MTB/RIF-generated  $C_T$ -values in different body compartments. In **Chapter 9** we address the important clinical question of whether, in the context of high routine empiric treatment rates, urine LAM strip testing can offer any incremental diagnostic utility.

In **Chapter 10** I summarise the findings of the studies and overall conclusions from across all the studies are presented. The implications of this research to the TB

diagnostic field in general are discussed and the priorities for future research are identified.

### **Coherence of thesis**

The coherence of this PhD thesis is underpinned by four common themes. Firstly, the research has the unifying theme of presenting novel diagnostic approaches to the same clinical challenge (SN- and SSTB) in the same geographical setting. Secondly, the work presents an evolution of investigation into the use of i) SI and ii) the novel MTB/RIF and LAM diagnostic tools from simple diagnostic accuracy measurements to impact assessments on treatment outcomes and the evaluation of incremental diagnostic utility in routine clinical settings. Thirdly, I (Jonathan Peter) was the first author on all ten of the included manuscripts and was the lead investigator in all of the studies (with the exception of the original prospective patient cohort, used for the local evaluation of MTB/RIF performance, which was lead by my supervisor, Professor Keertan Dheda). Finally, all these studies have been undertaken with the guidance and supervision of Professor Keertan Dheda during my time working at the UCT Lung Infection and Immunity unit over the past four years.

## References

1. Dheda K, Davids V, Lenders L, Roberts T, Meldau R, Ling D, Brunet L, van Zyl Smit R, Peter J, Green C, Badri M, Sechi L, Sharma S, Hoelscher M, Dawson R, Whitelaw A, Blackburn J, Pai M, Zumla A. Clinical utility of a commercial lam-elisa assay for tb diagnosis in Hiv-infected patients using urine and sputum samples. *PLoS One* 2010;5:e9848.
2. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, Milovic A, Jones M, O'Brien SM, Persing DH, Ruesch-Gerdes S, Gotuzzo E, Rodrigues C, Alland D, Perkins MD. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med* 2010;363:1005-1015.
3. Wirth T, Hildebrand F, Allix-Beguec C, Wolbeling F, Kubica T, Kremer K, van Soolingen D, Rusch-Gerdes S, Locht C, Brisse S, Meyer A, Supply P, Niemann S. Origin, spread and demography of the mycobacterium tuberculosis complex. *PLoS Pathog* 2008;4:e1000160.
4. WHO. Who report 2011: Global tuberculosis control. Geneva: World Health Organisation; 2011 No. ISBN 978 92 4 156438 0.
5. Abdool Karim SS, Churchyard GJ, Abdool Karim Q, Lawn SD. Hiv infection and tuberculosis in South Africa: An urgent need to escalate the public health response. *Lancet* 2009;374:921-933.
6. Lawn SD, Zumla AI. Tuberculosis. *Lancet* 2011;378:57-72.
7. Dowdy DW, Chaisson RE, Moulton LH, Dorman SE. The potential impact of enhanced diagnostic techniques for tuberculosis driven by Hiv: A mathematical model. *AIDS* 2006;20:751-762.

8. Keeler E, Perkins MD, Small P, Hanson C, Reed S, Cunningham J, Aledort JE, Hillborne L, Rafael ME, Girosi F, Dye C. Reducing the global burden of tuberculosis: The contribution of improved diagnostics. *Nature* 2006;444 Suppl 1:49-57.
9. Peter J, van Zyl Smit R, Denkinger C, Pai M. Diagnosis of tuberculosis: State-of-the-art. In: Lange C, Migliori GB, editors. *European monograph series: Tuberculosis*: European Respiratory Society; 2012. p. 124-143.
10. Getahun H, Harrington M, O'Brien R, Nunn P. Diagnosis of smear-negative pulmonary tuberculosis in people with Hiv infection or aids in resource-constrained settings: Informing urgent policy changes. *Lancet* 2007;369:2042-2049.
11. Lawn SD, Brooks SV, Kranzer K, Nicol MP, Whitelaw A, Vogt M, Bekker LG, Wood R. Screening for Hiv-associated tuberculosis and rifampicin resistance before antiretroviral therapy using the xpert mtb/rif assay: A prospective study. *PLoS Med* 2011;8:e1001067.
12. Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, Kraus S, Binder A, Meldau R, Hardy A, Dheda K. Diagnostic accuracy of a urine lipoarabinomannan strip-test for tb detection in Hiv-infected hospitalised patients. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* 2012;40:1211-1220.
13. Colebunders R, Bastian I. A review of the diagnosis and treatment of smear-negative pulmonary tuberculosis. *Int J Tuberc Lung Dis* 2000;4:97-107.
14. Banda H, Kang'ombe C, Harries AD, Nyangulu DS, Whitty CJ, Wirima JJ, Salaniponi FM, Maher D, Nunn P. Mortality rates and recurrent rates of tuberculosis

in patients with smear-negative pulmonary tuberculosis and tuberculous pleural effusion who have completed treatment. *Int J Tuberc Lung Dis* 2000;4:968-974.

15. Cohen T, Murray M, Wallengren K, Alvarez GG, Samuel EY, Wilson D. The prevalence and drug sensitivity of tuberculosis among patients dying in hospital in kwazulu-natal, South Africa: A postmortem study. *PLoS Med* 2010;7:e1000296.

16. Holtz TH, Kabera G, Mthiyane T, Zingoni T, Nadesan S, Ross D, Allen J, Chideya S, Sunpath H, Rustomjee R. Use of a who-recommended algorithm to reduce mortality in seriously ill patients with Hiv infection and smear-negative pulmonary tuberculosis in South Africa: An observational cohort study. *The Lancet infectious diseases* 2011;11:533-540.

17. Pai NP, Vadnais C, Denkinger C, Engel N, Pai M. Point-of-care testing for infectious diseases: Diversity, complexity, and barriers in low- and middle-income countries. *PLoS medicine* 2012;9:e1001306.

18. Khan MS, Dar O, Sismanidis C, Shah K, Godfrey-Faussett P. Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: A pragmatic randomised controlled trial. *Lancet* 2007;369:1955-1960.

## **Chapter 2.**

### **Background and Literature review**

**Modified and updated sections of the following three published reviews on TB diagnostics are included:**

#### **i. Diagnosis of tuberculosis: state-of-the-art**

Peter, J.G., van Zyl-Smit, R.N., Denkinger, C.M. and Pai M

Published: *European Respiratory Monograph* 2012; 58: 124-143

#### **ii. The progression of TB diagnosis in the HIV era: from microscopes to molecules and back to the bedside**

Peter, J. G. and Theron, G.

Published: *Continuing Medical Education* October 2011 Vol 29 No 10

#### **iii. Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments**

Peter, J.G., Green, C., Hoelscher, M., Mwaba, P., Zumla, A. and Dheda K

Published: *Current Opinions in Pulmonary Medicine* 2010, 16:262-270

## Summary

Rapid, affordable and accurate TB diagnosis is key to effective patient management and global TB control. Effective clinical screening and optimized sample acquisition methods remain the first steps in the diagnostic process. Smear microscopy, despite optimization, remains widely used even though sensitivity is poor. Mycobacterial liquid culture is accurate, but poorly accessible and delayed. The use of novel molecular tools such as Xpert MTB/RIF or Genotype MTBDRplus assays, which offer superior diagnostic accuracy and decreased time-to-diagnosis for drug-sensitive and/or resistant TB, is increasing following WHO endorsement and in some countries national rollout is underway. In contrast, both serology (antibody-detection tests) and IGRAs have been found to offer little diagnostic utility for active TB diagnosis and have been discouraged by WHO. Other novel, simple technologies, such as the POC urine LAM strip test and the visually read LAMP NAAT, although of uncertain and restricted clinical utility, highlight the progression toward an inexpensive, instrument-free, laboratory-free POC diagnostic technology, which will be essential to tackle TB in the future.

## Introduction

Affordable, accurate and rapid diagnosis followed by effective therapy is the cornerstone of tuberculosis (TB) control. TB is curable in over 95% of cases but the same diagnostic standard remains elusive to many who need it most. Multiple social, host and pathogen factors, as depicted in Figure 1, intersect to produce and worsen TB diagnostic delay or failure. Distinct diagnostic priorities and requirements exist today in several different high burden settings. Availability of healthcare resources, integrity of diagnostic services, and HIV co-infection impact on the efficacy of traditional diagnostic tools. Smear microscopy, more than a century old, is simply no longer adequate to meet these challenges. Fortunately, thanks to renewed global awareness, financial investment and international collaboration, several new diagnostic options have been developed. Old tools continue to be optimized and several new tools are now commercially available and being up-scaled by national TB programmes.

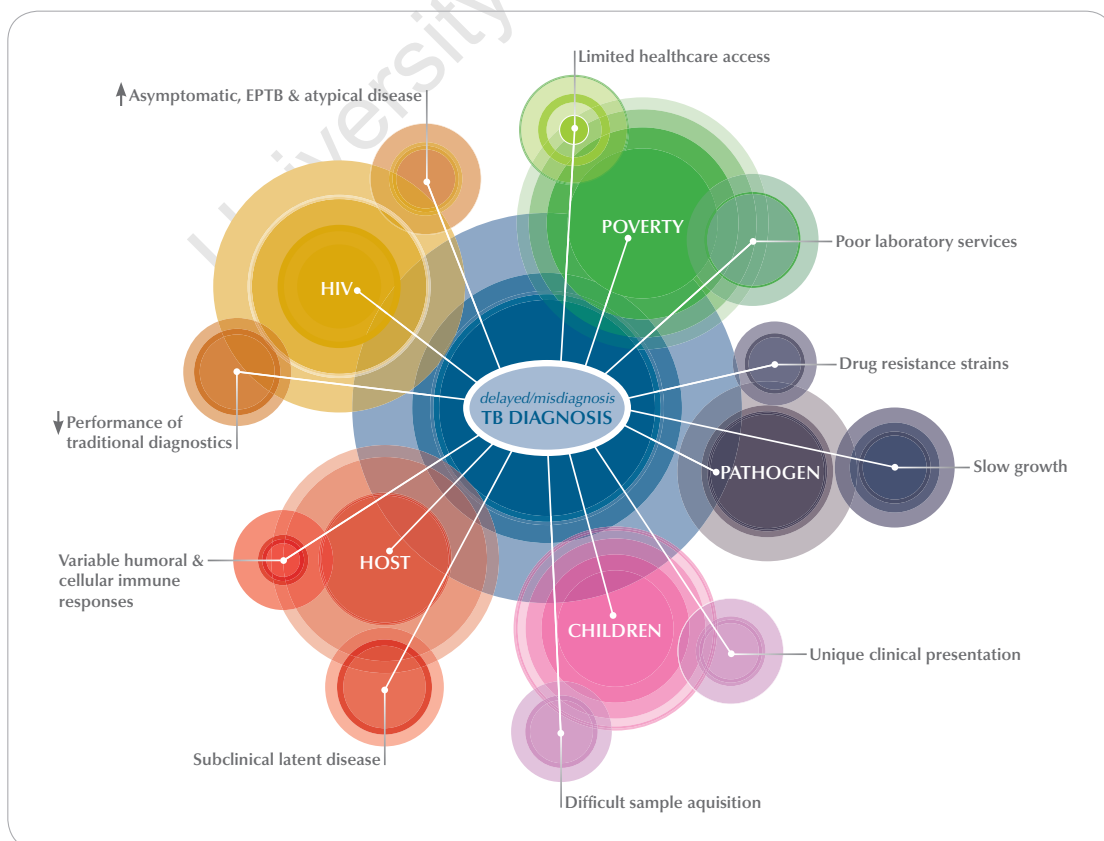


Figure 1. The multiple inter-related factors driving misdiagnosis or delayed diagnosis of TB. The diameters of the larger circles indicate the relative impacts on delayed/misdiagnosis of TB and the overlapping circles indicate relatedness. EPTB: extrapulmonary TB, DR: drug-resistant [Overlapping circles represent interconnectedness but are not based on comparative data].

The TB diagnostic landscape and need continues to be shaped by the intersection of available diagnostic technologies, national healthcare resources and differing National TB programs as well as TB epidemiology. In high burden countries (HBC) the priority remains the diagnosis of active, predominantly pulmonary TB and drug susceptibility testing. In countries with the dual epidemic of HIV this extends particularly to smear-negative, sputum-scarce and/or extrapulmonary forms of TB (EPTB), the burden of which was outlined in Chapter 1. In HBC, healthcare resources and laboratory infrastructure are limited, affected populations are generally poor, frequently rural or marginalized with restricted access to health services. Thus, the need for affordable, laboratory-independent, point-of-care technologies is critical. In contrast, the diagnostic priority of low burden countries, in addition to active TB, is focused on detecting latent TB infection and preventing the development of active disease.

This background review will discuss the optimization of old tools and the development and integration of new diagnostic technologies for active TB. In particular, diagnostic test performance characteristics in SNTB and optimal settings for use will be discussed. Additionally, two *focus in* sections (Sample acquisition methods and urine-based TB diagnosis) are provided to provide in-depth background relevant to the studies of the PhD thesis on sputum induction and the Urine LAM strip test. Finally key research 'gaps' and ongoing unmet diagnostic needs will be emphasized.

## **Review methodology**

I searched electronic databases for both relevant recent review articles and primary research studies for the latest data pertaining to TB diagnosis. I searched Pubmed, Embase and Web of Science databases at different time-points during my PhD thesis preparation, with the most recent searches conducted at the end of 2012. Relevant search terms for each of the specific diagnostic topics covered were used, including: i) For the focus in section on sample acquisition methods, I included “tuberculosis”, “Mycobacterium tuberculosis” AND “sputum induction”, “induced sputum”, “sputum sampling”, “sputum sample acquisition”, “assisted sputum sampling”, “diagnostic sample acquisition”, “pulmonary sampling”, “bronchoalveolar lavage”, “bronchoscopy”, “healthcare-worker assisted sampling”, “provider-training”, “sputum submission instruction”, “chest physiotherapy”, “gastric washing”, and “nasopharyngeal aspirate”. These searches were narrowed down in pubmed with the use of the “diagnosis[broad]” filter; and ii) For the focus in section on urine-based TB diagnosis, I included “tuberculosis”, “Mycobacterium tuberculosis” AND “urine”, “urinary”, “trans-renal DNA”, “lipoarabinomannan”, “LAM”, and “urine-based diagnosis”. In addition, I identified additional studies by contacting experts in the field, and by searching reference lists from primary studies, review articles, and textbook chapters. Finally, the published reviews forming the basis for this background review chapter had additional relevant studies that I had missed added by other co-authors and editors.

Table 1. Commercially available diagnostics for active TB and drug-susceptibility testing [References up to end of 2012]

Test type or platform	Description of test	Current validated commercial versions	High/Low sensitivity (Expected) need to cite references for accuracy estimates	High/Low specificity (Expected)	WHO endorsement	Comments
<b>A. Detection of Active TB</b>						
Fluorescence microscopy using LED	Auramine-O stained smear read by fluorescent microscopy using LED light source	Primo Star iLED™ (Carl Zeiss, Germany) Lumin™(LW scientific, GA, USA) and others	56-80% (concentrated, direct samples compared to culture) (23-26)	92-98% (compared to culture) (24-26)	Yes	Approximately 10% greater sensitivity vs. ZN light microscopy No reduction in performance in HIV co-infection (23) WHO endorsed
Semi-automated, nonintegrated NAAT	Amplification and detection of mycobacterial rRNA or DNA direct from clinical samples	Amplified MTD* (GenProbe, San Diego, CA, USA) Probe Tec ET (BD, Franklin Lakes, NJ, USA); Cobas Taqman MTB (Roche Molecular diagnostics, CA, USA)	36-100% (pooled approx. 66-96%) (27-29)	54-100% (pooled 85-98%) (27-29)	No	Sensitivity in smear-negative patients 50-80% (28, 29) Open system at risk for DNA contamination and Specificity affected by laboratory quality control
Simplified, manual NAAT	Isothermal amplification with visual readout to detect mycobacterial DNA direct from clinical samples	Eiken LAMP* (Eiken, Japan)	Overall: 83% Smear positive patients: >95% Smear negative: 41-52% (30) (IUTLD conf. proceedings)	>97% (30)	No	Not yet approved by WHO expert group. Awaiting results of demonstration study
Serological (antibody) detection test	Immunological test: detection of antibodies to TB antigens by ELISA or rapid lateral flow format	Although several assays are on the market, no currently available test has been validated and proven to be clinically useful	0-100% (31) Anda-TB IgG: Pooled estimates in smear positive: 76% smear negative: 59%	31-100% (31) Anda-TB IgG Pooled estimate: 92%	No	WHO made negative recommendation in 2011
Antigen detection test	TB antigens detected by ELISA or lateral flow test format	Alere TB LAM ELISA (Waltham, MA, USA)	Overall: 18-59% HIV only: 20-67% (32)	88-100% (32)	No	Only offers clinical utility in HIV-infected patients with advanced immunosuppression Validation of lateral flow strip test (Determine TB*) ongoing
<b>B. Detection of Active TB and DST</b>						
Fully-automated,	Fully automated, self contained	Xpert MTB/RIF* (Cepheid,	MTB detection:	MTB detection:	Yes	WHO strong recommendation

integrated NAAT	platform integrating sputum processing, mycobacterial DNA extraction and amplification	CA, USA)	Overall: 90% Smear positive patients: 94-100% Smear negative patients: 46-83% (pooled sens: 75% Rifampicin resistance: 98-99% (33-36)	>98% Rifampicin resistance: >98% (33)		for frontline TB diagnosis in HIV-infected and MDR TB suspects
Automated liquid culture with indirect DST	Automated system of for mycobacterial liquid culture and subsequent DST	BACTEC MGIT® 960 (BD Diagnostics, MD, USA), BacT/ALERT® 3D (bioMerieux, France)	100% in smear positive cases, >75% in smear negative cases	>99%	Yes	Approximately 10% higher diagnostic yields compared to solid media culture Contamination can be 10-20% in laboratories with poor quality assurance Indirect DST can take 3-4 months to provide results
Non-automated liquid culture with direct DST	Simplified systems for mycobacterial liquid culture with reduced laboratory equipment for MTB detection and direct DST	TB MODS Kit® (Hardy Diagnostics, Santa Maria, CA, USA)	96% (compared to traditional automated liquid culture) (37)	96% (compared to traditional automated liquid culture) (37)	Yes	Low equipment requirements offset by high labour needs Direct DST provides results in 10-14 days
Phage-based detection	Bacteriophage viruses infect and detect the presence of MTB	FASTPlaque™ (Ipswich, Suffolk, UK)	81-100% (38, 39)	73-100% (38)	No	3-36% indeterminate rate limits use; Not endorsed by WHO
<b>C. DST and/or speciation</b>						
Manual amplification and hybridization (Line probe assay)	NAAT with hybridization of amplified product to strip test allowing for identification of MTB and common mutations causing resistance to rifampicin and INH	GenoTypeMTBDRplus® (Hain LifeScience) INNO-LiPARif.TB® (Innogenetics)	For Rifampicin resistance: >98% For INH: >84% (40, 41)	For rifampicin resistance: >98%, for INH: >99%	Yes	Provides DST results on culture isolates and smear positive clinical specimens in 1-2 days Undergoing widespread upscaling in NTPs with the help of the EXPAND TB program
Rapid speciation assay	Rapid immunochromatographic (lateral flow) test for identification MTB complex in culture isolates	Capilla TB-Neo® (Tauns), TBc ID® (BD diagnostics, MD, USA), SD Bioline Ag MPT64 Rapid® (Standard Diagnostics Inc, Korea)	>98%	>99%	Yes	Simple, reliable tests for use especially in settings with high rates of NTMs

LED: light emitting diode; NAAT: Nucleic acid amplification test; MTB: *Mycobacterium tuberculosis*; FIND: Foundation of innovative and novel diagnostics; LAM:

lipoarabinomannan; MDR: Multi-drug resistant; DST: Drug susceptibility testing; NTM: non-tuberculous mycobacteria, NTP=National TB Program

## **Diagnosis of active pulmonary TB and drug-susceptibility testing**

This section considers current diagnostic practice and tools for the diagnosis of active pulmonary TB and drug-susceptibility testing. The optimisation of traditional diagnostic methodology and tools is discussed together with the development and appropriate use of new diagnostic technologies. Table 1 describes current, commercially available TB diagnostics for active TB and drug susceptibility stratified by the test's ability to provide diagnostic and DST alone or combined. In addition, to better contextualize available old and new tests, figure 2 visually contrasts the test performance and time-to-result of commercially available diagnostics for active TB and DST.

### **“Improving the old”**

#### *Clinical case definitions and symptom screening*

Long before Koch's discovery of the tuberculosis bacilli, physicians had recognized and clinically described TB disease. Clinical presentations of TB disease are numerous; TB can infect almost every organ and mimic numerous other human diseases (1); and our evolving understanding of subclinical TB infection suggest a complex host-pathogen interaction and a changing spectrum of sub-clinical infection (2). These factors, combined with the absence of accurate or affordable diagnostic tools especially in HBC, means that clinical screening, diagnosis and case definitions continue to guide treatment decisions (including for isoniazid preventive therapy), form part of composite diagnostic reference standards (especially in children and HIV-infected patients), and direct suspected TB patients for further diagnostic testing. Symptom screening and clinical case definitions to guide the use of isoniazid preventive therapy (IPT) and empiric TB treatment in persons living with HIV and smear-negative TB respectively are areas of

particular clinical and public health importance. A recent meta-analysis, including 9626 people, found the combination of any one of: current cough (any duration), fever, night sweats or weight loss to have a sensitivity and specificity of 79% and 50%, respectively (3). At TB prevalence rate of 5 and 20%, the absence of any of these symptoms could reliably exclude active TB in 98% and 90% of patients, respectively (3). The addition of chest x-ray (CXR) to symptom screening improved sensitivity by 12%, but reduced specificity by 11% (3). The WHO have used these findings in the recently published intensified case finding (ICF) and IPT guideline (WHO) (4). Given the well-established burden of asymptomatic active TB amongst HIV infected patients with advancing immunosuppression (5), the limitations of symptom screening methods are clear and application should be cautious, although their use may remain necessary for public health TB screening, IPT and ICF strategies.

In 2006, the WHO developed the smear-negative TB guidelines, with expanded case definitions for both pulmonary and extrapulmonary TB (6). These guidelines are widely available and help guide treatment decisions in many settings, yet few studies have evaluated their diagnostic accuracy and impact. In ambulatory patients attending an outpatient clinic, the diagnostic accuracy was modest, with a sensitivity and specificity of 80% and 44% respectively (7). Despite modest diagnostic performance, another South African study in hospitalized patients however, suggested that the strict use of clinical guidelines and case definitions might potentially impact patient mortality and hospital length of stay (8). Studies are ongoing to improve the use of clinical treatment algorithms, integrate them with the newer diagnostic tools such as MTB/RIF, consider their impact (and the use of non-algorithm guided empiric treatment) on the evaluation of novel diagnostics and evaluate their impact on patient important outcomes such as mortality in HIV infected patients with advanced immunosuppression.

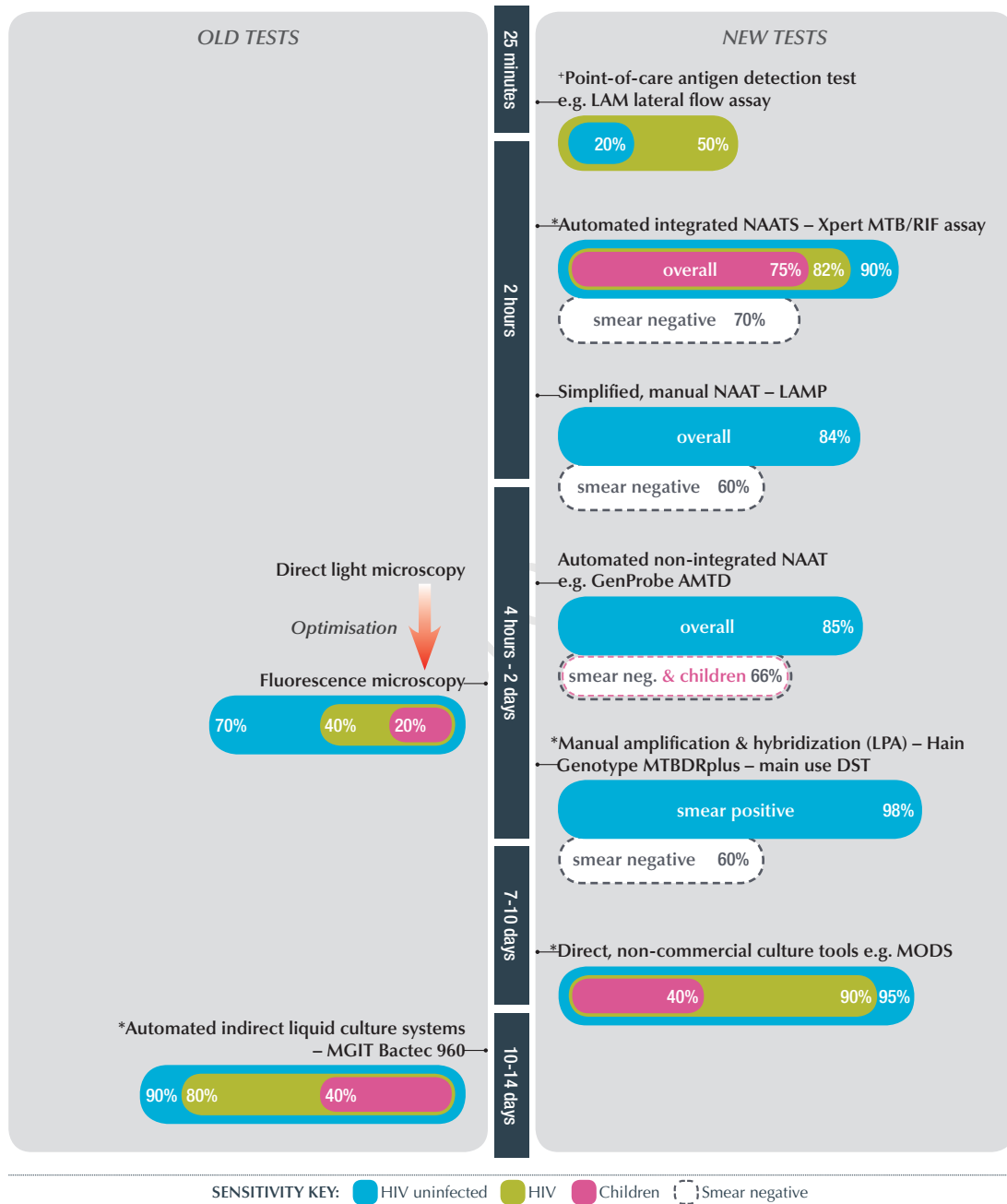


Figure 2. Comparison of the sensitivity (indicated by percentages) and time-to-diagnosis for active pulmonary TB and drug susceptibility diagnostic tools indicating areas of reduced performance in HIV TB co-infected patients and children. Only tests commercially available and with a specificity >95% for the diagnosis of active TB are

included. LAM: Lipoarabinomannan; NAATs: Nucleic-acid amplification tests; LAMP: Loop isothermal amplification PCR; DST: drug-susceptibility testing; MODS: Microscopic Observed Drug Susceptibility; MGIT: Microscopic Growth-in-tube

### *Chest radiology*

Radiology is widely utilized in both high and low burden settings for both TB screening in asymptomatic patients and the diagnosis of active disease (9). Used alone, chest x-ray (CXR) has only moderate specificity and in settings of high HIV-prevalence also moderate sensitivity (10), with between 10-71% of HIV co-infected TB patients having an entirely normal CXR despite culture-positive disease (5, 11-13). However, when CXR is used in conjunction with other simple diagnostic tools such as symptom screening and/or smear microscopy, it can offer both diagnostic utility and cost-efficacy particularly for ruling-out active TB disease (14-16). Two South African studies found that the combination of symptoms with/without sputum smear microscopy followed by CXR offered a negative predictive value of more than 95% for active TB in a high burden setting (17-19). Unfortunately, optimal CXR utility requires interpretation by trained, skilled observers, which are not always available. Inter-observer variability of CXR has been shown to be poor irrespective of reader skill (10, 20, 21). To overcome this drawback, several radiological scoring systems, such as the Chest Radiograph Reading system (CRRS), have been developed to improve inter-observer variability (13). Furthermore, automated computer systems to interpret and report digital CXRs are currently in-development (22).

## 1<sup>st</sup> FOCUS IN SECTION:

### **The acquisition of samples for diagnostic testing**

Definitive TB diagnosis relies on the demonstration of *M. tuberculosis* bacilli, TB-specific antigens or genetic material (PCR). For this, an appropriate and sufficient biological sample is essential. However, attaining an adequate sample can be challenging and can result in either no confirmed TB diagnosis or a substantial delay in diagnosis. In HBC, especially in countries with high HIV co-infection and childhood TB prevalence, a significant proportion of patients are unable to spontaneously produce sputum (42, 43). Patients are unfamiliar with techniques to optimally expectorate sputum or feel uncomfortable about producing sputum (44, 45). It has been very elegantly shown that sputum production is not equivalent in men and women, with reduced smear microscopy sensitivity being seen in females from a number of countries (44, 45). A number of strategies and techniques have been evaluated to improve sputum expectoration, to induce sputum or attain an alternative pulmonary sample suitable for laboratory TB diagnosis. These strategies and techniques for sample acquisition and their advantages and disadvantages are outlined in Table 2.

#### *Sputum induction*

Of the techniques outlined in Table 2 below, sputum induction is emerging as the preferred sampling strategy for use in resource limited settings. Sputum induction has been found to be safe and tolerable in children, adolescents and adults irrespective of HIV co-infection (46, 47). It can be performed with minimal infrastructure and low-cost infection control booths have been developed for use in urban and rural, and even mobile clinic settings. The infection control booth pictured below was built for this project at Langa primary care clinic.



Photographs of facilities used for sputum induction in resource-limited settings including i) a community clinic low-cost infection control sputum booth, ii) a rural tobacco plantation clinic with battery-powered nebuliser, and iii) a mobile tent infection control booth (starting top left and moving clockwise)

Table 2. Strategies/Techniques for the improved acquisition of pulmonary samples for TB diagnosis

Technique	Diagnostic performance ranges†	Advantages	Disadvantages	Knowledge 'gaps'
<b>Expectorated sputum assistance techniques</b>				
Provider-training and observed sputum collection	(Malawian SN-TB suspects (48)) 39/46 (85%) definite-TB cases detected	Minimal staff training requirements Inexpensive Widely applicable in all settings	Time-consuming Infection control risk Not applicable for children	Programmatic research on implementation, uptake and efficacy
Sputum submission instructions/training	(Pakistan women study (45)) ↑smear positive case detection ↓spot-sputum saliva submission ↑women returning with sputum (Indonesia men/women study (49)) 15% higher case detection	Minimal staff training requirements Inexpensive Widely applicable in all settings No infection control risk	Time-consuming Not applicable for children	Programmatic research on implementation, uptake and efficacy
<b>Sputum induction techniques</b>				
Physical manoeuvres (e.g. chest physiotherapy)	Diagnostic yield (TB culture) Adults: 5-26% (48, 50) Smear sensitivity (TB culture reference std) Adults: 50-53% (48, 50)	Safe procedure Minimal training/no equipment requirements	Low diagnostic yield High infection risk for health worker Currently restricted to hospitals with trained physiotherapists	No studies in primary care settings and children Few comparison studies with other methods of sputum induction
Ultrasonic nebulisation‡	Diagnostic yield (TB culture) Adults: 8-34% (48, 51) Children: 10-30% (46, 52) Smear sensitivity (TB culture reference std) Adults: 37-78% (51, 53) Children§: 20-57% (54, 55)	Safe procedure Non-invasive Feasible in resource-poor settings Good yield in adults and children, Simpler performance in HIV-infected and uninfected patients	Equipment and consumable costs High infection risk for health worker Currently restricted to district hospitals with infection control facilities	Few studies in primary care settings No studies of impact on patient-important outcomes, positioning in diagnostic algorithms, and use of novel diagnostic tools on induced sputum samples
Other devices e.g. vibration tool such as Lung flute	43% smear microscopy sensitivity (56) (small study of 15 patients)	Safe procedure Non-invasive Disposable/self-explanatory equipment decreases infection risk	High costs and waste of device Feasibility in different settings Infection risk Tools still in development	Prospective studies required in clinically appropriate settings

<b>Alternative respiratory sample acquisition techniques</b>				
Gastric washings <sup>‡</sup>	Diagnostic yield (TB culture) Adults: 11-30% (57, 58) Children: 5-17% (46, 59) Smear sensitivity (TB culture reference std) Adults: 30-37% (48, 58) Children: 18-53%(46, 59)	Safe and effective procedure especially for children Minimal infection risk	Invasive procedure Requires fasting Sample collection advised on 3 consecutive days Not feasible in many public health facilities Currently restricted to district-level hospital settings	Programmatic studies of yields from resource-poor settings
Nasopharyngeal aspirate <sup>‡</sup>	Diagnostic yield (TB culture) Children: 7-9% (59, 60) Smear sensitivity (TB culture reference std) Children: 58-71% (59, 61)	Safe Feasible across settings Useful for diagnosis of other respiratory pathogens e.g. viruses especially in children	Low diagnostic yields for TB Infection risk Currently restricted to district-level hospital	Programmatic studies of diagnostic utility in routine clinic settings
Bronchoscopy <sup>‡</sup>	Diagnostic yield (TB culture) Adults: 9-46% (48, 62)  Smear sensitivity (TB culture reference std) Adults: 27-63% (63, 64)	Equivalent diagnostic yield to other sample acquisition methods but allows direct visualization of respiratory tract ± biopsy Allows diagnosis of other respiratory pathogens e.g. pneumocystis	Invasive Requires specialized equipment and staff Restricted to tertiary and district hospitals Expensive Infection risk for health worker	Studies on the performance of novel diagnostics using bronchoalveolar lavage fluid to diagnosis or exclude active TB

<sup>†</sup>Diagnostic performance characteristics are from prospective studies predominantly in high-burden countries. Wide variation and heterogeneity predominantly accounted for by differences in included study populations (e.g. smear-negative TB suspects vs. smear-negative TB suspects with chest radiography suggestive of TB) and background TB prevalence. In induced sputum studies using ultrasonic nebulisation, adult studies include only smear-negative/sputum scarce TB suspects, while child studies are of TB suspects without prior diagnostic testing. Performance outcomes for novel diagnostics applied to acquired pulmonary samples are not included.

<sup>‡</sup>Comparative studies of some/all of these sample acquisition techniques have been done for adults (48, 50, 58, 63) and children (46, 59, 65). For both adults and children sputum induction using ultrasonic nebulisation is equivalent or superior to the more invasive techniques of gastric washing and bronchoscopy.

<sup>§</sup>A recent study of 452 children with suspected TB found diagnostic yield of induced sputum to be 15% and the sensitivity of the novel diagnostic, Xpert MTB/RIF to be 83% (58/70) (66).

Induced sputum, using ultrasonic nebulisation of hypertonic saline, has been found to have higher culture-based TB diagnostic yields than nasopharyngeal and gastric aspirates and be equivalent to bronchoscopy (67, 68). Consequently, sputum induction is now the preferred sampling method in children and has shown promising results in HIV-infected patients (69, 70). However, data are limited for adults with SNTB and no direct comparative data quantifying differences in performance between important patient sub-groups is available to inform service needs (11, 48, 51, 71). Further research is needed into the performance characteristics in primary care clinic settings, the impact on patient-important outcomes and the positioning and cost-efficacy of integrating sputum induction into NTP diagnostic algorithms.

University of Cape Town

## **“Improving the old” (continued)**

### *Smear microscopy*

Although widely used, the sensitivity of smear microscopy is highly variable, ranging between 20- 80% (72), performing poorest in HIV-infected patients (10) and children (69) . Additionally, smear microscopy relies on well-trained microscopists and sensitivities between field and reference laboratories can vary by as much as 28% (35)

The most important developments in optimizing smear microscopy and associated WHO policy changes are outlined in Figure 3A. In addition, several innovative approaches to further improve smear microscopy are under development including improved concentration techniques using nanobeads, fluorescence microscopy with molecular probes (FISH), automated and computer-assisted smear reading technologies, and use of mobile phones for microscopy (73, 74).

### *The expansion and development of culture based techniques*

*Mtb* (*Mycobacterium tuberculosis*) culture remains the clinical and research diagnostic gold standard for all forms of active TB. Figure 3B outlines the progress and associated WHO policy changes for mycobacterial culture techniques for both diagnosis and drug-susceptibility testing (DST). Traditional solid culture methods are tedious, time-consuming and have limited clinical impact. Automated liquid culture systems, with approximately 10% higher yields and a decreased time-to-diagnosis (75) have largely replace solid culture. However, automated liquid cultures are expensive, prone to contamination, and require considerable laboratory infrastructure and expertise. Thus, despite WHO endorsement in 2007, they remain inaccessible to populations where they are most needed. In 2009, the WHO endorsed the use of alternative, simpler, less expensive non-commercial culture and drug-susceptibility testing technologies, as an interim measure while capacity for genotypic testing is scaled-up (76). Details of the

microscopic-observed drug susceptibility (MODS) method are shown in table 1, while other endorsed non-commercial methods include the colorimetric redox-indicator (CRI) and the nitrate reductase assay (NRA). Phage-based methods have not been WHO-endorsed due to insufficient evidence, variable specificity and high rates of invalid results (77). Under controlled laboratory conditions, these non-commercial culture methods are inexpensive and can provide culture and DST results in 7-14 days (76, 78). Lack of standardization and local variations in methodology remain programmatic concerns and have thus far limited scale-up.

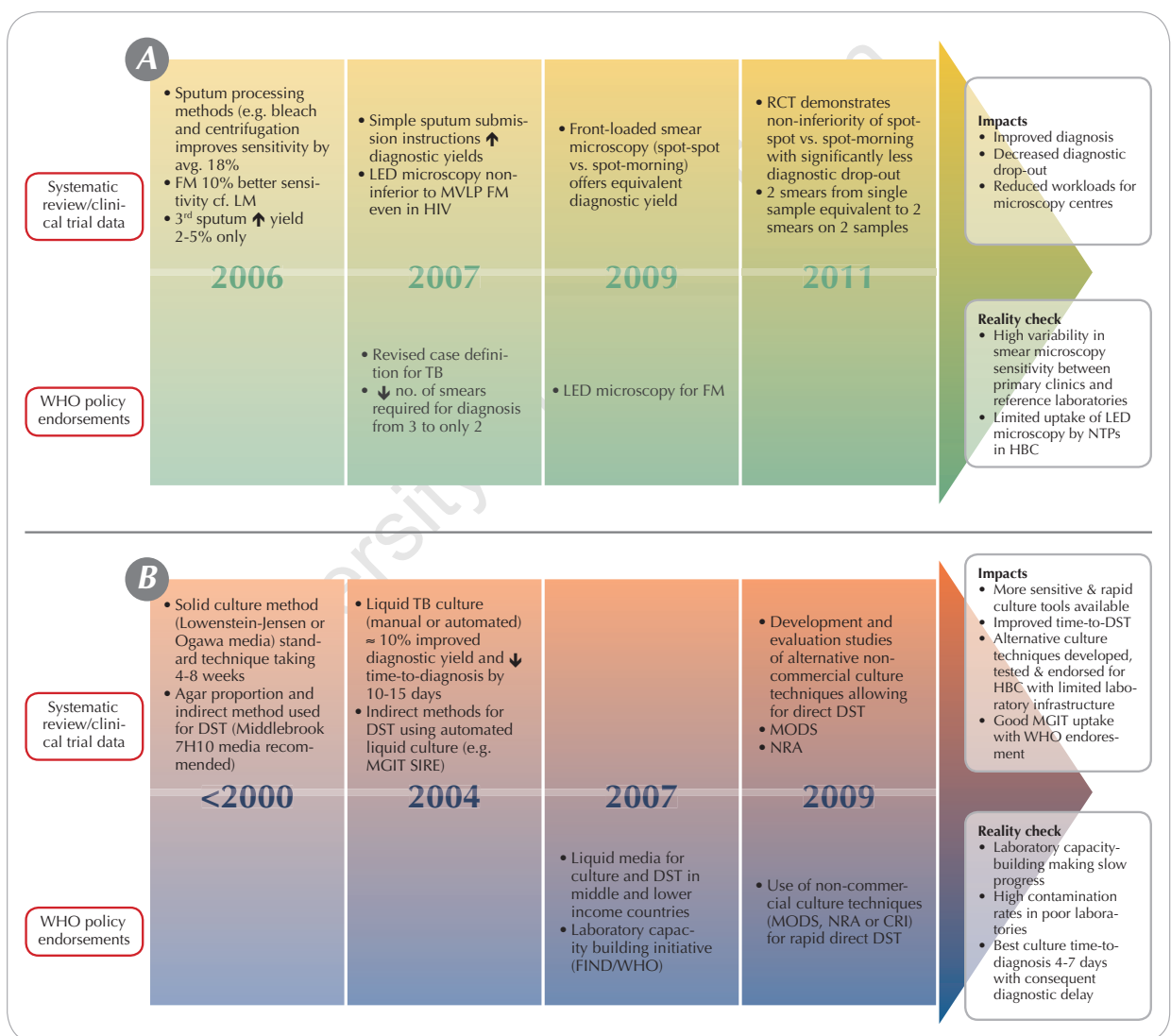


Figure 3. Progress in optimising and streamlining A) sputum smear microscopy and B) tuberculosis (TB) culture. FM: Fluorescence microscopy; LM: Light microscopy; LED: Light

emitting diode; MVLP: Mercury vapour lamp; NTP: National TB programmes; HBC: High burden countries; DST: drug susceptibility testing; MGIT SIRE: Microscopic growth-in-tube sensitivities isoniazid, rifampicin and ethambutol; MODS: Microscopic observed drug susceptibility; NRA: nitrate reductase assay; CRI: colorimetric redox-indicator

## **“Ushering in and tailoring the new”**

### *Non-integrated, semi-automated nucleic acid amplification tests*

Conventional, non-integrated NAATs (see table 1 for details and examples) have been found to offer high specificity (85-98%) and sensitivity for smear-positive TB (~96%), but poorer sensitivity (~60%) and specificity for smear-negative TB (27-29). Compared to smear and culture, these assays are expensive, requiring specialised laboratory infrastructure and expertise; while, being open systems, are at risk for cross-contamination in settings with sub-optimal laboratory quality. These factors have limited their widespread uptake in high-burden resource-limited settings. Simplified, manual NAATs, such as LAMP (Eiken, Japan) using isothermal amplification and a visual readout have been developed as more affordable options where laboratory infrastructure is limited (30). Early evaluation suggested similar performance to other commercial NAATs with a sensitivity of ~40% in smear-negative TB (30). The Foundation of Innovative and Novel Diagnostics (FIND) is currently conducting large-scale evaluation and demonstration studies of LAMP, with promising preliminary results (79). However, despite the assays' simplicity, the risk of cross-contamination during manual DNA extraction and need for laboratory training and technical skill remain and may prevent widespread application.

### *Integrated, fully automated NAAT: Xpert MTB/RIF*

In December 2010, the WHO announced the endorsement of the novel MTB/RIF assay (Cepheid, CA, USA) (80). MTB/RIF is a fully automated and integrated DNA extraction and amplification system, thereby addressing many limitations of existing commercial NAATs (81). Furthermore, MTB/RIF has the potential to be performed in decentralized locations outside of reference laboratories by staff with minimal laboratory training (1-2 days).

To date, the MTB/RIF assay has undergone evaluation in sputum samples from more than 11000 patients in 19 countries (33, 35, 36, 43, 66, 82-84), although these studies have performed MTB/RIF in a laboratory, rather than at point-of-care (POC). A meta-analysis of these 18 published studies show a sensitivity and specificity of a single sputum-based MTB/RIF for culture positive TB of 90.4% (89.2-91.4) and 98.4% (98.0-98.7) respectively and 75% for smear-negative culture positive (S-C+) PTB (33). Performing a second and third MTB/RIF increases sensitivity by approximately 13% and 5% respectively (36), while indeterminate rates of only 1-3%, decreasing to <1% after repeat testing, have been found across settings (35, 36).

Importantly, the use of MTB/RIF decreased the mean time-to-treatment initiation amongst smear-negative culture positive TB patients from 56 to 5 days – similar to that of smear-positive patients (35). For the detection of rifampicin resistance, the meta-analysis data shows a sensitivity and specificity of 94.1% and 97.0% respectively (33). On this evidence base, WHO has made a strong recommendation for the use of frontline MTB/RIF in all patients with suspected drug-resistant TB and/or co-infected with HIV, and a conditional recommendation, in acknowledgment of resource implications, for the use of MTB/RIF as a follow-on test to microscopy in settings where MDR-TB or HIV is of lesser concern, especially for smear-negative TB (80).

Undoubtedly, a number of unanswered questions and concerns around the use of the MTB/RIF assay remain. Firstly, although specificity for detecting rifampicin resistance remains >98%, studies continue to find false positive rifampicin resistance results (85). In areas of low MDR-TB prevalence, given the significant decrease in positive predictive value associated with small reductions in specificity, a large number of false positive rifampicin resistance results may occur with widespread routine use. Despite the development of updated versions of both the GeneXpert cartridge and software to further improve assay specificity, this remains an important concern. Secondly, given the reduced ability of a single MTB/RIF test to ‘rule-out’ TB in HIV-infected patients (34), the role of additional MTB/RIF tests and alternative investigations in HIV-infected patients with ongoing symptoms needs to be better defined. Thirdly, given that MTB/RIF detects both viable and non-viable *Mtb*, the interpretation of a positive MTB/RIF in patients not responding to TB therapy and the use, if any, of MTB/RIF for treatment monitoring requires urgent clarification. Finally, a number of operational challenges and research questions associated with the national and international scale-up of MTB/RIF and other new TB diagnostic technologies remain and are outlined in Figure 4.

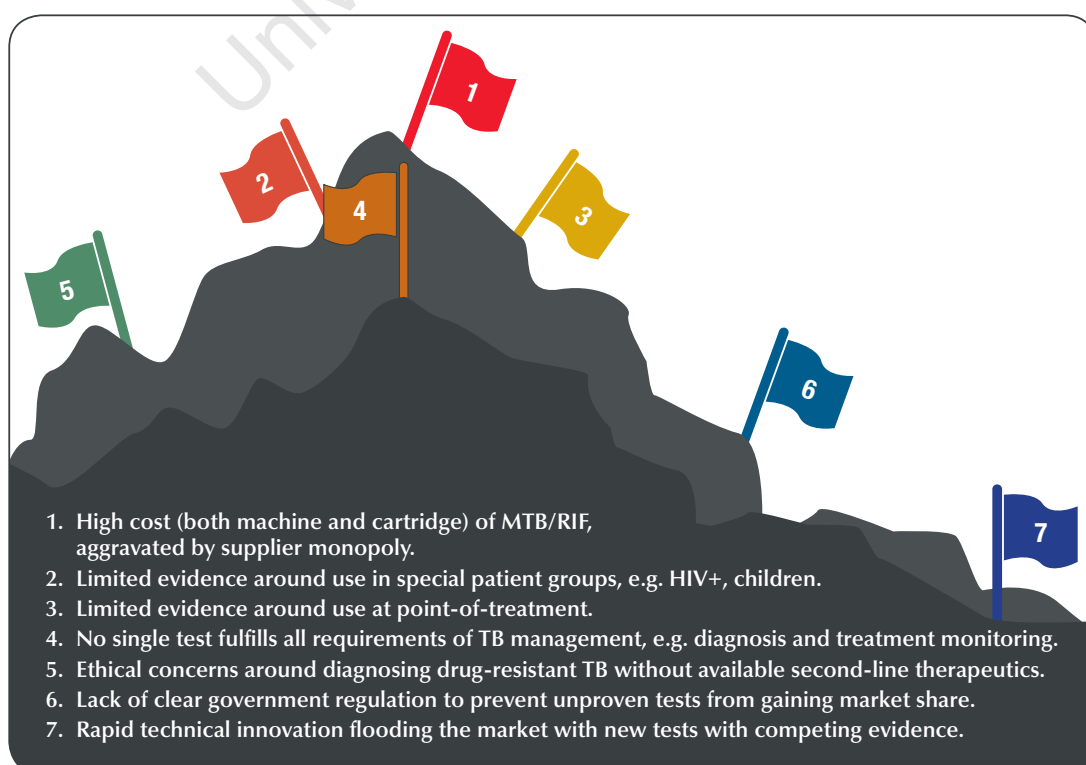


Figure 4. The challenges of up-scale and implementation of MTB/RIF and other novel TB diagnostic technologies

#### *Line probe assays*

As a rapid alternative to phenotypic DST, specialised NAATs, using manual amplification and hybridization techniques, known as line probe assays (LPA) (table 1) offer *Mtb* speciation and genotypic DST with results in 1-2 days. LPAs received WHO endorsement in 2008 as the test of choice for rapid genotypic rifampicin and isoniazid DST (86). LPAs offer sensitivities >98% for rifampicin resistance, but only ~85% for INH resistance due to the presence of resistance coding mutations outside the regions of the *inhA* and *katG* genes detected by the assays (40, 87). LPAs are now routinely available and being scaled-up in certain national TB programmes e.g. South Africa and India by the EXPAND-TB project for MDR-TB suspects with smear- or culture-positive samples. Recently, a Genotype MTBDRsl (Hain Lifescience GmbH, Nehren, Germany) has been developed and is now available for rapid genotypic 2<sup>nd</sup> line DST and XDR-TB diagnosis. Sensitivities vary across initial studies depending on the specific drug tested (88-90) and clinical utility is restricted to rapidly 'ruling-in' XDR-TB at this stage.

#### *Immunodiagnosis for active TB*

Numerous serological tests to detect TB-specific antibodies are available in many developing countries (91). Updated meta-analyses show current serological assays are of no clinical value with high variability in both sensitivity and specificity (31), and poor cost-effectiveness (92). Despite the demonstrated lack of either accuracy or cost-efficacy, these tests continue to be sold in 17 of 22 high burden settings (91). In India alone, an estimated US\$15 million per annum is spent on performing serological tests for TB in the private sector (91). In 2011, in response to this, the WHO issued a negative policy advising against the use of any of the numerous available

blood serological assays for the diagnosis of TB (93), and countries such as India have banned the clinical use of these tests. However, this policy does not discourage ongoing research into serological tests for TB diagnosis.

Blood-based interferon-gamma release assays (IGRAs) in high burden settings have also been extensively evaluated and found to offer little, if any, clinical utility as a frontline diagnostic for active TB in either HIV-infected or uninfected patients (19, 94-96). A recent WHO policy (2011) discourages use of IGRAs for active TB diagnosis in low and middle income countries (97). TST remains a useful tool for the diagnosis of active TB in young children and IGRAs offer equivalent, but not superior performance (98, 99). It is clear that immunodiagnosis is not a substitute for molecular or microbiological site-of-disease diagnosis, although its use for investigating latent TB infection remains important and is discussed below.

## 2<sup>nd</sup> FOCUS IN SECTION

### Urine-based diagnosis of active TB

The use of biological samples other than sputum such as urine, exhaled breath, and blood for the diagnosis of TB is particularly appealing given the challenge of attaining sputum or equivalent pulmonary samples for diagnostic testing. Of these, urine is particularly appealing because of its availability, ease of access (easily obtained from children and adults), processing and storage, and the low infection risk to health-care workers during sample collection.

#### *Conventional mycobacterial smear and culture using urine specimens*

As early as the 1960's it was postulated that *M. tuberculosis* bacilli could be found in the urine of patients with active pulmonary tuberculosis, and for many years it was part of the microbiological work-up of patients with suspected TB. Currently, it is not recommended for the routine diagnosis of pulmonary TB as two large retrospective reviews indicated that the yield of urine smear and culture was less than 2% (100, 101). However, in HIV-infected patients with advanced immunosuppression the yield of urine culture is improved (77% in one HIV-infected cohort (102)). However, even in this subgroup the sensitivity of urine smear-microscopy remains poor as does the incremental yield (~ 5%) of urine *M. tuberculosis* culture over sputum or lymph node smear-microscopy or culture (103). Thus, conventional TB diagnostics applied to urine specimens has limited clinical usefulness.

## Urinary LAM for the diagnosis of TB

The detection of mycobacterial antigens in urine for diagnostic purposes has been suggested since the 1930s (104). The 17.5 kD glycolipid, lipoarabinomannan (LAM), found in the outer cell-wall of mycobacterial species has been investigated as a diagnostic antigen. LAM, an immunogenic virulence factor, is released from metabolically active or degrading bacterial cells (105) during TB infection (106). LAM is heat stable, filtered by the kidney and detectable in the urine (107-110) (Figure 5). Published data suggest that LAM is specific for mycobacterial species (111, 112). However, LAM-like glycolipids are found in several species of fungi and bacteria including *Corynebacterium diphtheria* (113). We have described that anti-LAM polyclonal antibodies cross-react with various Actinobacteria, including different strains of *Nocardia* and *Streptomyces*, and *C. albicans*, and *T. paurometabolum* found in normal mouth flora (114).

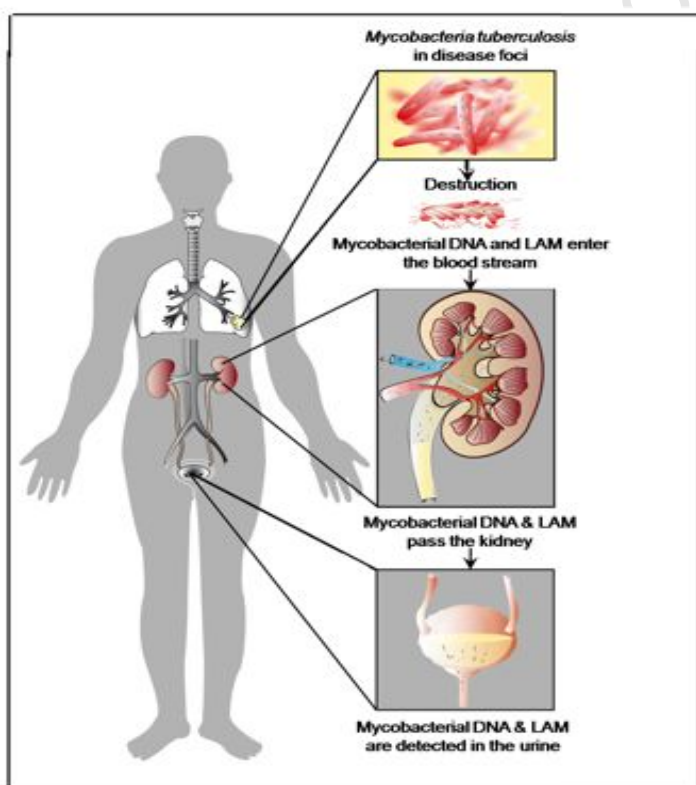


Figure 5. Annotated diagram illustrating the passage of mycobacterial DNA and LAM antigen from infection site to urine

Hamasur and co-workers developed an ELISA and dipstick test for the detection of LAM antigen in urine (107). These early developmental

studies were followed by the creation of a 1<sup>st</sup> generation pre-commercial prototype [MTB LAM ELISA® (Chemogen, Portland, USA)], which was clinically trialled in 2005 in Tanzania (109). A second generation prototype using a different batch of polyclonal antibody was then developed. This was recently superseded by a commercially available prototype, the Clearview® TB ELISA (Alere, USA), which uses the same polyclonal antibody used in the 2<sup>nd</sup> generation Chemogen assay. Integrated commercial dipstick prototypes have been developed using the same polyclonal antibody and are currently undergoing validation studies as potential user-friendly and rapid point-of-care diagnostic tests (including the validation study presented in this thesis).

*Clinical utility of urine LAM in unselected patients with suspected TB from a primary care setting*

Tessema et al, reported a sensitivity and specificity (95% CI) of 74.0% (67-80) and 86.9% (84-89), respectively (108) in an Ethiopian cohort where the HIV status was not determined.

Boehme *et al*, using the 1<sup>st</sup> generation MTB LAM ELISA® prototype (Chemogen, Portland, USA) reported a sensitivity and specificity of 80.3% and 99%, respectively, in a Tanzanian cohort with a HIV seroprevalence of 69% (109). Unfortunately, the initial promise of these early studies was not sustained. Five published studies now show consistently poor overall test performance with sensitivities ranging between 13 to 51% (114-118)(Table 3). These disappointing results suggest that there is little clinical utility for the urine LAM assay in unselected outpatients with suspected TB. Nevertheless, across studies in high TB and HIV prevalent settings test sensitivity was higher (21 to 62%) in the HIV-infected patients (114, 116, 117). However, which HIV-infected outpatients are most likely to benefit from the urine LAM assay is unclear because of the lack of stratification by CD4 counts in previous studies (115-117).

### *Urinary LAM and worsening immunosuppression*

Interestingly, a crude inverse relationship can be demonstrated between improved urine LAM sensitivity and declining CD4 count. Thus, when several settings are considered, the test performs best in patients with advanced immunosuppression i.e. CD4 < 200 cells/mm<sup>3</sup> (114, 119-121). Why is there an increase in LAM positivity in those with advanced immunosuppression? The reasons are not entirely clear. HIV-infected patients with advanced immunosuppression have more severe disease and a likely higher antigen burden (122). Boehme *et al.* found a correlation between degrees of smear positivity (1+, 2+ etc.) and LAM positivity, suggesting that burden of disease is related to the degree of urine LAM positivity (109). We hypothesise that changes in glomerular filtration secondary to HIV-related podocyte dysfunction (123), commoner in advanced HIV, may account for the observed increase in urinary LAM as may the presence of mycobacteriuria (124). In addition, a decrease in the complexing of LAM to circulating antibodies in the dysregulated immune response associated with HIV-infection may also play a role (125).

Table 3. Performance of standardized urine lipoarabinomannan assays in different clinical subgroups of patients with suspected tuberculosis stratified by HIV status and, where available, CD4 T cell count

Study	Clinical subgroup and country	Evaluable pts tested/ Total number	HIV (%pos.)	Overall sensitivity(CI), HIV pos. pts, HIV pos. pts with CD4< 200µl/ml	Overall specificity(CI), HIV pos. pts, HIV pos. pts with	Overall PPV(CI), HIV pos. pts, HIV pos. pts with CD4< 200µl/ml	Overall NPV(CI), HIV pos. pts, HIV pos. pts with CD4< 200µl/ml	Sensitivity (CI) in smear(S) neg. culture(C) pos. TB cases (n)
Boehme et al. (109)	Outpatient clinic and healthy USA/local controls (Tanzania)	235/333	69.0%	OVERALL 80 HIV+ 81 CD4<200 N/R	99 (N/R) N/R N/R	100 (N/R) N/R N/R	80 (N/R) N/R N/R	(n=50) 76 (N/R) N/R N/R
Daley et al.# (115)	Outpatient clinic (India)	200/200	8.5%	OVERALL 18(9-33) HIV+ 20(1-77) CD4<200 N/R	88(81-92) 83(51-97) N/R	30(15-50) 33(2-88) N/R	79(72-84) 71(42-90) N/R	(n=12) 25 (4-64) N/R N/R
Reither et al.# (117)	Outpatient clinic (Tanzania)	151/291	59.1%	OVERALL 51(38-63) HIV+ 62(47-75) CD4<200 N/R	88(79-94) 84(68-94) N/R	78(63-89) 84(68-94) N/R	68(58-77) 62(47-75) N/R	(n=21) 38 (N/R) 42 (N/R) N/R
Mutetwa et al.# (116)	Outpatient clinic (Zimbabwe)	261/397	77%	OVERALL 44(36-54) HIV+ 52(43-62) CD4<200 N/R	89(81-94) 86(73-93) N/R	84(74-91) 84(73-92) N/R	54(46-61) 57(48-66) N/R	(n=40) 28 (13-43) N/R N/R
Dheda et al.*1 (114)	Outpatient clinic (South Africa)	427/500	31%	OVERALL 13(8-19) HIV+ 21(11-35) CD4<200 37(19-59)	99(97-100) 100(91-100) 100(84-100)	94(74-99) 100(70-100) 78(45-94)	59(53-64) 55(43-65) 76(63-86)	(n=70) 3 (N/R) 18 (N/R) 25 (N/R)
Lawn et al.# (126)	ARV clinic, asymptomatic HIV patients (South Africa)	235/235	100%	OVERALL 38 (27-51) HIV + 38 (27-51) CD4 50-100 41(22-64) CD4< 50 67(44-84)	100 (N/R) 100 (N/R) N/R N/R	100(82-100) 100(82-100) N/R N/R	83(77-88) 83(77-88) N/R N/R	(n=50) 36 <sup>+</sup> (N/R) 36 <sup>+</sup> (N/R) 35 <sup>+</sup> (N/R)
Shah et al. *1(119)	In-patients (South Africa)	315/499	85%	OVERALL 59 (52-66) HIV+ 67 (59-74) CD4 50-100 71 (51-87) CD4 <50 85 (73-93)	96(91-99) 94(87-98) N/R N/R	73(65-80) 75(67-82) N/R N/R	34(29-39) 30(25-36) N/R N/R	(n=111) 56 (N/R) N/R N/R N/R
Gounder et al.*1 (118)	Outpatient clinic (South Africa)	422/423	100%	OVERALL N/A HIV+ 32 (16-52) CD4 <200 35 (16-57) CD4 <50 56 (21-86)	N/A 98(96-99) 96 (92-99) 90 (76-97)	N/A 53 (28-77) 57 (29-82) 56 (21-86)	N/A 95 (93-97) 92 (87-95) 90 (76-97)	N/A (n=16) 16 (4-40) N/R N/R
Talbot et al.*1(121)	In-patients (Tanzania)	212/278	100%	OVERALL N/A HIV+ 65 (53-76) CD4 50-200 67 (43-85) CD4 <50 77 (58-89)	N/A 86 (79-91) 94 (79-99) 72 (56-84)	N/A 69 (56-80) 88 (60-98) 67 (49-81)	N/A 84 (76-89) 82 (66-92) 82 (65-92)	N/A (n=24) N/R N/R N/R

# Boehme et al used the 1<sup>st</sup> version of the MTB LAM ELISA prototype (Chemogen, Portland, USA) and the other four studies (#) used a 2<sup>nd</sup> version (different polyclonal antibody) of the MTB LAM ELISA prototype (Chemogen, Portland, USA); \*1 these studies used the commercially available Clearview® TB ELISA assay (Inverness Medical Innovations, Waltham, MA). N/A: Not applicable; N/R = not reported in study manuscript

### *Clinical utility of urine LAM in hospitalized TB suspects with advanced immunosuppression*

The improved diagnostic performance with worsening immunosuppression meant that urine LAM testing may offer the greatest diagnostic utility in hospitalised TB HIV co-infected patients with more advanced immunosuppression and a greater severity of TB disease. Shah *et al* in a nested cohort study of 499 South African in-patients with suspected TB (85% HIV-infected) found the overall sensitivity (95% CI) of the Clearview® LAM ELISA amongst confirmed TB cases to be 59% (52-66) (119). This sensitivity improved to 71% (51-87%) in patients with a CD4 50-100 cells/mm<sup>3</sup>, and 85% (73-93) in patients with CD4 < 50 cells/mm<sup>3</sup> (119). It is significant that the urine LAM sensitivity in smear negative TB patient group was 56% (119). A more recent hospital study in Tanzanian in-patients found a similar sensitivity in patients with severe immunosuppression (121).

### *Clinical utility of urine LAM for TB screening in HAART-naïve ambulatory HIV positive patients*

HAART has substantially reduced the risk of HIV-infected individuals acquiring TB (127, 128) and mortality in patients with active TB (128, 129). However, HIV-infected patients commencing HAART, because of several considerations including unmasking TB-IRIS, must first be screened for active TB (130, 131). In Africa ~ 25% of patients without a known TB diagnosis and referred for HAART were found to have culture-positive TB when investigated further; the majority of patients had smear-negative TB (126, 132, 133). Lawn *et al*, in a South African study evaluating 235 HAART-naïve patients with a median CD4 count of 125 cells/ml demonstrated a urine LAM sensitivity (95% CI) of 33% (22-46) using unconcentrated urine (126). In this study standard diagnostic tools such as TB symptom screening and chest x-ray showed a sensitivity of 64%, but a specificity of just 39%, whilst smear-microscopy showed a sensitivity of only 14% (126). Combined sensitivity of urine

LAM and sputum smear-microscopy was 45% but this increased to 67% if confined to patients with a CD4 count <50 cells/mm<sup>3</sup>. Given the limited data it is not possible to make clinical recommendations but these preliminary results suggest that urine LAM may be a useful adjunctive diagnostic tool for TB screening in ARV roll-out clinics in high TB burden countries where patients often present with advanced immunosuppression. Further studies are now needed to clarify these findings.

#### *Clinical utility of LAM ELISA using biological samples other than urine*

Although the majority of studies have evaluated the diagnostic utility of the LAM ELISA using urine in different clinic settings, the detection of LAM antigen has now been explored in several studies using various biological samples including serum (134, 135), sputum (136), cerebrospinal fluid (137) and pleural fluid (138). In serum, detection of LAM is complicated by immune complex formation (134), while utility in both sputum and induced sputum detection is complicated by cross-reactivity with mouth flora e.g. *Candida* spp. (114, 139). In pleural fluid it has been shown to be diagnostically unhelpful (138) but a recent proof-of-concept study showed that, when used on cerebrospinal fluid, it may be a promising adjunct tool for the diagnosis of TB meningitis (137).

#### *Concerns about urine LAM specificity*

A major concern when considering the application of urine LAM to routine clinical practice is the poor test specificity (83 to 89%) found in four clinical studies with varying HIV prevalence (115-117, 121). This contrasted with the excellent test specificity (99 to 100%) found in three other recent trials (114, 119, 120). The poorer specificity in some studies may possibly be explained by undiagnosed occult TB disease, detection of antigen from latently-infected subjects (unlikely given the suboptimal test sensitivity), contamination of the sample by environmental mycobacteria or other bacteria (particularly if the sample

was collected at home and brought to the clinic), or contamination by non-bacterial species e.g. *Candida*, which is prevalent in HIV-infected populations. These possibilities are tenable given our findings that several organisms, including *Candida* species, cross-react with the polyclonal LAM antibodies within the assay (114). These issues are important to consider in future studies.

#### *Point-of-care immunochromographic strip format of the urine LAM assay*

Although the diagnostic accuracy of urinary LAM means that clinical utility is limited to TB HIV co-infected patients with advanced immunosuppression, both the particular diagnostic difficulties with this vulnerable patient group together with the recent development of a point-of care dipstick urine LAM prototype (Determine TB<sup>®</sup> LAM Ag strip test, Alere, USA) makes this test ideally suited for use in resource-limited high HIV prevalent setting. The urine LAM strip test is an instrument-free, laboratory-free, affordable (<US\$3.5) POC test producing results within 25 minutes and using an easily obtainable urine sample with low infectious risk (140). The evaluation of this novel point-of-care test is one of the main focuses of this PhD work and further test details are described in subsequent chapters (manuscripts).

#### ***Urine PCR for the diagnosis of tuberculosis***

Alongside the cell wall components of *M. tuberculosis*, nucleic acid fragments have also provided a target for detection in the diagnosis of pulmonary and extrapulmonary TB. Initial reports of detecting *M. tuberculosis* DNA in the urine of patients without detecting whole *M. tuberculosis* bacilli first reached the literature in 1997 (141), and since then a

number of studies have reported MTB DNA detection without genito-urinary TB disease (reviewed in (142)). PCR remains a modestly sensitive and highly specific tool that can accurately quantify *M. tuberculosis* load and make a speedy diagnosis. These factors have led to the development of commercial platforms such as the self contained GeneXpert (Cepheid) for near beside diagnosis of tuberculosis and drug resistance from sputum. Yet urine PCR remains un-utilised. The one unifying characteristic of this field is the variability of reported sensitivity, specificity and methodology. Unlike LAM, no commercial assay has been developed specifically for the detection of urinary trans-renal DNA and all methods remain in-house. Direct comparisons of published methods are missing, and when such a study was conducted by our group the results were not comparable to the previous reports (143).

Considering the extent of confusion, are there any hard and fast rules for applying PCR technology to urine in a diagnostic setting? Certainly, a greater understanding of the phenomenon leading to the presence of DNA in urine (144) (Figure 5) has provided consensus that, for PCR diagnosis of non-genito-urinary disease, urine should not be centrifuged prior to nucleic acid extraction (145). The sensitivity of PCR detection is also improved by reducing the size of the target (146) though careful assay design is essential to preserve specificity. HIV infection does seem to correlate with an increase in detection of MTB DNA in urine, though the reason for this has not been determined or accurately correlated with severity of either disease.

At present, urine PCR for the routine diagnosis of pulmonary TB is not recommended, but this approach should not be completely discarded. Where current diagnostic gold standards are sub-standard in many patient groups, urine could be an adjunct sample. A recent study of pulmonary TB in India demonstrated that of 11 bacteriologically negative

but urine PCR positive cases, 9 (81.8 %) responded to anti-TB therapy (ATT) (147). This study circumvented any temporal changes in *M. tuberculosis* DNA in urine by pooling specimens collected over a 3 day period, but did extract from the pellet of centrifuged samples and amplify a large target (786 bp) possibly reducing detection sensitivity. We are currently working to determine the impact of temporally focusing collection and ATT on *M. tuberculosis* DNA in urine.

Optimising PCR sensitivity and specificity by addressing specimen collection, storage (145) and extraction are the first steps to producing a urine PCR that delivers for pulmonary TB diagnosis. The next steps are to determine the relevance of *M. tuberculosis* DNA in urine to TB disease. It is clear that latency and persistent mycobacteria (148) are challenges for the current diagnostic repertoire. Techniques do exist, such as quantitative PCR, which provide extremely accurate measurements of *M. tuberculosis* DNA, but the correlation between the amount of fragmented *M. tuberculosis* DNA in urine and TB disease needs to be determined. The diagnostic capability of any new test should also not be considered in isolation of the point of application. The highest burden of TB still remains in the most resource poor countries, where the application of advanced molecular techniques requires specific considerations (149). A point-of-care test for trans-renal *M. tuberculosis* DNA would be a huge step forward for molecular based diagnosis from urine.

### ***Additional experimental approaches for urine-based diagnostics***

Given the advantages of using urine as a diagnostic sample, together with the fact that urine likely contains a number of host and organism-related proteins and metabolites, the identification of signature molecules for TB diagnosis in the urine hold much promise and warrant ongoing interest. We, and others, are currently using proteomic and metabolomic

technologies to screen for urinary peptides and metabolites in an attempt to uncover candidate molecules and/or characteristic TB-specific protein signatures. Napolitano *et al*, using these techniques, identified four TB specific antigens, including ornithine carboamyltransferase, which may be useful for the diagnosis and monitoring of PTB (150). Additionally, investigators are evaluating monoclonal antibodies against known TB antigens, including LAM and the RD-1 antigens, for the diagnosis of TB using urine samples. Other urine-based approaches including use of an electronic nose (E-nose) (151) and gas chromatography-mass spectrometry to characterize volatile organic compounds specific to TB are also being investigated. Urine PCR, if found to be useful, will likely be improved by simpler and user-friendlier NAAT platforms such as loop mediated isothermal amplification (LAMP) (30).

### **Towards point-of-care technology for active TB**

Despite the advancement in the molecular diagnosis for TB and drug-resistance, the need for a simple, instrument-free, laboratory-free POC test continues to be articulated by both research groups and civil societies (152-154). Required minimum specifications for the ideal POC TB test have been defined by a number of groups, recently published (155) and are highlighted in textbox 1 below. Mathematical models suggest a huge potential impact of POC TB diagnosis on both case detection and overall TB incidence (156, 157).

### **Textbox 1.** Optimal characteristics of a future point-of-care TB (POC) diagnostic test

▫

#### **Test performance characteristics**

- Sensitive (>95% for smear-positive and >60% smear-negative culture-positive samples) and specific (>95% compared to culture)
- User-friendly (ease of sample collection and processing, minimal technical and training requirements, results available preferably within a single patient-healthcare contact that enables treatment initiation if available and if appropriate)
- Robust (shelf-life >24 months, stable in high temperature and humidity, can be battery powered, easy waste disposal)

#### **Other important considerations**

- Affordable (cost < US\$10/test) and accessible in high burden countries
- Targeted to one or more specific healthcare levels e.g. home, community, clinic, peripheral lab, hospital; each level has specific user and device requirements

Some commercially available novel diagnostics already come close to meeting these requirements and may offer important POC utility. The MTB/RIF assay, easily meets the specifications for diagnostic test accuracy (Sensitivity: >95% for S+C+ and 60-80% for S-C+; Specificity: >95%) and time-to-result (<3 hours), but falls short as an ideal decentralized POC test because of its cost and the specialised equipment needs. Feasibility and impact studies of point-of-treatment, clinic-based MTB/RIF are nearing completion (<http://www.clinicaltrials.gov/ct2/show/NCT01554384>) and will provide insights on its' POC utility. For TB HIV co-infected patients with advanced immunosuppression, the Determine® TB LAM Ag strip test (Alere, USA) offers POC potential as discussed above and in the further chapters of this thesis. Neither of these tests provide the ideal POC TB test, yet they indicate that the development of such a test may be within reach (152).

The ongoing progression towards POC TB diagnosis is graphically outlined in Figure 4. The diagnostic pipeline has many promising molecular point-of-care tests and platforms under development. Hand-held or portable platforms using DNA chips and/or disposable

cartridges are being evaluated for POC, simplified NAATs (153), while technologies to transition ELISA assays into simplified lateral flow POC test formats are well established and being increasingly exploited. Although currently commercial serological tests are inaccurate for TB diagnosis, the detection of individual or combinations of TB-specific antibodies, antigens and or immune-markers using lateral flow assays or microfluidic technologies still seems most likely to provide a field-friendly POC tool (153, 158). In addition, both platforms seem to be evolving toward simultaneous detection and diagnosis of different infectious disease. Finally, electronic nose technology allowing analysis of breath condensates and the detection of distinct profiles of volatile organic compounds offers another possibility for POC TB diagnosis (159, 160).



Figure 6. Current progress and future evolution of TB diagnosis from smear microscopy to molecular methods and onwards towards simple, affordable point-of-care test formats

## **Unmet needs and research priorities**

Social, environmental, host and pathogen-specific factors continue to create distinct diagnostic challenges and settings (Figure 1), both at individual patient- and public-health levels. No single test has yet, or perhaps ever-will meet all diagnostic requirements across resource-, healthcare- and clinical-settings. Integration of old and novel technologies and continued tailoring of technology to individual, high and low burden, local and national settings is essential to optimize TB diagnosis. Table 4 highlights many of the current unmet diagnostic needs and research 'gaps'. Ongoing basic and clinical research, as well as increased operational research, will be required to address these gaps. In particular, research moving beyond the simple assessment of diagnostic accuracy towards impact evaluations of novel tools and integrated algorithms on important patient and public health outcomes such as morbidity, mortality, case detection rates and/or default rates, and hospital length of stay (161) are required to best develop and guide policy. Recently, Cobelens et al. proposed a new phased evaluation pathway for TB diagnostics (Figure 7).

## **Conclusion**

The armamentarium of diagnostics tests for tuberculosis has never been greater. Nevertheless, many diagnostic challenges on both an individual patient-level, such as for smear-negative or sputum scarce TB, EPTB, TB HIV co-infection and childhood TB, as well as on larger public-health level remain sub-optimally addressed. Novel diagnostic approaches and technologies are urgently required to meet these challenges.

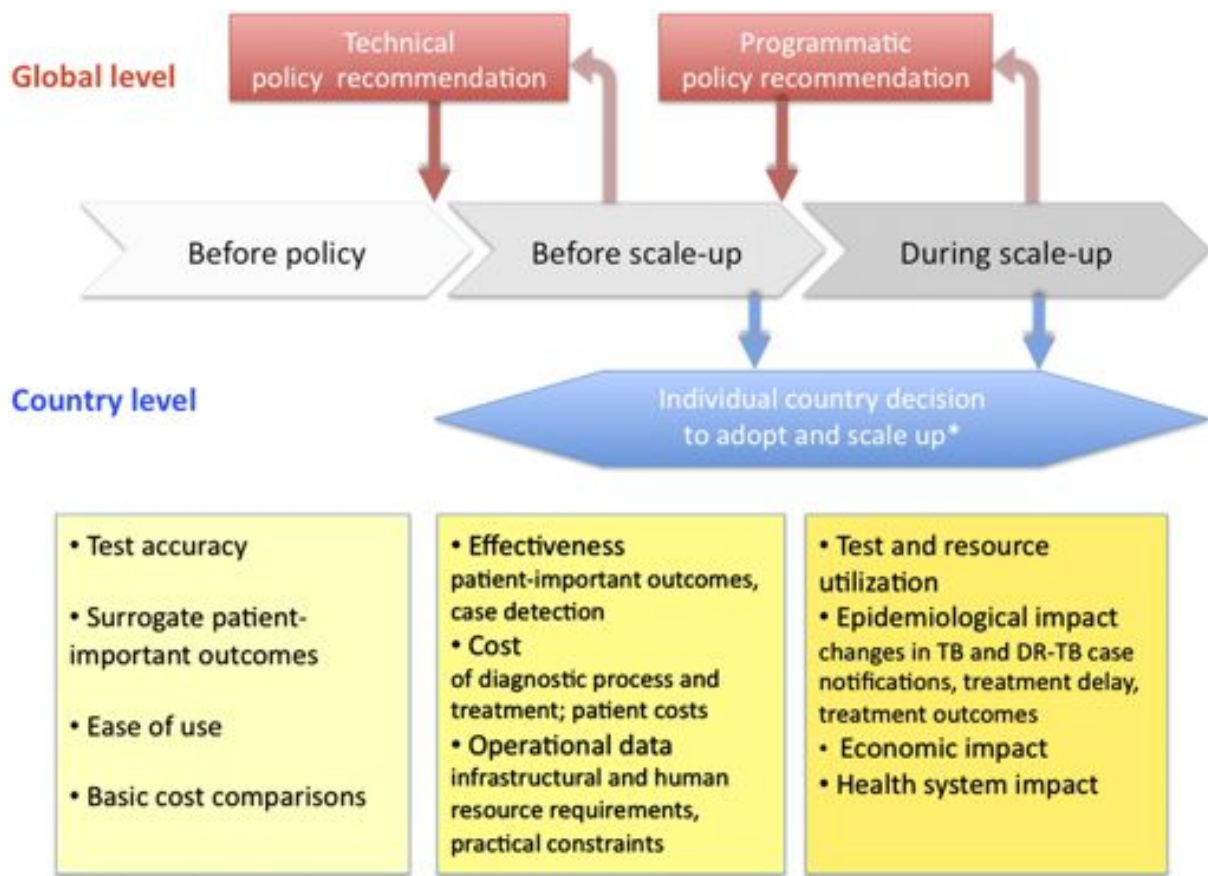


Figure 7. Proposed new value chain for phased evaluation of TB diagnostics, from accuracy to impact assessment (161), reproduced with permission from Journal of Infectious Disease

Table 4. Important current unmet TB diagnostic needs and research ‘gaps’

<b>Research ‘gap’ and/or unmet diagnostic need</b>	<b>Rationale for need and/or research question(s)</b>
1. Development of a simple, affordable, field-friendly POC for active TB using sputum samples	High burden countries, severely limited resources Poor laboratory infrastructure and technical skills Patients have difficulty accessing health-services and default prior to diagnosis
2. Impact evaluations of different simple and safe sample acquisition techniques e.g. sputum induction for sputum-scarce, smear-negative and childhood TB in primary care settings	Up to a 1/3 of patient in high HIV and TB prevalent settings are unable to produce sputum All TB diagnosis relies on an adequate sample Sputum induction is simple and feasible yet carries high infection risk and moderate cost
3. Impact evaluations of MTB/RIF at different healthcare levels; operational research and cost efficacy evaluations of MTB/RIF; Optimal positioning of MTB/RIF in diagnostic algorithms	Rapid WHO endorsement and plan for global implementation Benefits of a 2-hour test result may be lost if not used at point-of-treatment and rapidly available to patients Operational performance and actual cost efficacy unknown
4. Development of rapid, non-sputum based POC test for the diagnosis of EPTB and childhood TB	EPTB and children most often unable to produce sputum Biological samples such as urine readily available Certain forms of EPTB e.g. TB meningitis carry very high mortality and rapid diagnosis would save lives
5. Development of a rapid ‘rule-out’ test for TB HIV co-infection for use in high burden settings	TB can be atypical clinically in HIV co-infection but progresses rapidly with high mortality rate High TB drug-related morbidity in HIV-infected patients Other pathogens can mimic TB presentation and cause mortality if untreated
6. Further studies and impact evaluation of available POC urine LAM strip test for HIV-infected patients with advanced immunosuppression	First simple, affordable, rapid, non-sputum based TB diagnostic available Targets HIV co-infected patients with advanced immunosuppression and highest TB-related mortality Lack of clarity about test specificity, cut-point selection and test patient impact
7. Development of simple to perform, improved rapid molecular assays for first and second-line drug resistance	Growing epidemic of MDR and XDR-TB All phenotypic DST methods require at least 10-14 days to provide results
8. Predictive biomarker(s) to identify latently infected people likely to progress to active TB and who will benefit most from preventive therapy	IGRA and TST predict progression to active TB sub-optimally Isoniazid preventive therapy can cause significant individual morbidity and require large public health expenditure

## References

1. Davies PD, Pai M. The diagnosis and misdiagnosis of tuberculosis. *Int J Tuberc Lung Dis* 2008;12:1226-1234.
2. Barry CE, 3rd, Boshoff HI, Dartois V, Dick T, Ehrt S, Flynn J, Schnappinger D, Wilkinson RJ, Young D. The spectrum of latent tuberculosis: Rethinking the biology and intervention strategies. *Nature reviews Microbiology* 2009;7:845-855.
3. Getahun H, Kittikraisak W, Heilig CM, Corbett EL, Ayles H, Cain KP, Grant AD, Churchyard GJ, Kimerling M, Shah S, Lawn SD, Wood R, Maartens G, Granich R, Date AA, Varma JK. Development of a standardized screening rule for tuberculosis in people living with hiv in resource-constrained settings: Individual participant data meta-analysis of observational studies. *PLoS Med* 2011;8:e1000391.
4. WHO. Guidelines for intensified tuberculosis case-finding and isoniazid preventive therapy for people living with hiv in resource-constrained settings. Geneva; 2011.
5. Oni T, Burke R, Tsekela R, Bangani N, Seldon R, Gideon HP, Wood K, Wilkinson KA, Ottenhoff TH, Wilkinson RJ. High prevalence of subclinical tuberculosis in hiv-1-infected persons without advanced immunodeficiency: Implications for tb screening. *Thorax* 2011 Aug;66:669-673.
6. WHO. Improving the diagnosis and treatment of smear-negative pulmonary and extrapulmonary tuberculosis among adults and adolescents. Geneva, Switzerland; 2006.
7. Wilson D, Mbhele L, Badri M, Morroni C, Nachega J, Chaisson RE, Maartens G. Evaluation of the world health organization algorithm for the diagnosis of hiv-associated sputum smear-negative tuberculosis. *Int J Tuberc Lung Dis* 2011;15:919-924.

8. Holtz TH, Kabera G, Mthiyane T, Zingoni T, Nadesan S, Ross D, Allen J, Chideya S, Sunpath H, Rustomjee R. Use of a who-recommended algorithm to reduce mortality in seriously ill patients with hiv infection and smear-negative pulmonary tuberculosis in south africa: An observational cohort study. *The Lancet infectious diseases* 2011;11:533-540.
9. Bakari M, Arbeit RD, Mtei L, Lyimo J, Waddell R, Matee M, Cole BF, Tvaroha S, Horsburgh CR, Soini H, Pallangyo K, von Reyn CF. Basis for treatment of tuberculosis among hiv-infected patients in tanzania: The role of chest x-ray and sputum culture. *BMC Infect Dis* 2008;8:32.
10. Reid MJ, Shah NS. Approaches to tuberculosis screening and diagnosis in people with hiv in resource-limited settings. *Lancet Infect Dis* 2009;9:173-184.
11. Wilson D, Nachega J, Morroni C, Chaisson R, Maartens G. Diagnosing smear-negative tuberculosis using case definitions and treatment response in hiv-infected adults. *Int J Tuberc Lung Dis* 2006;10:31-38.
12. Yoo SD, Cattamanchi A, den Boon S, Worodria W, Kitembo H, Huang L, Davis JL. Clinical significance of normal chest radiographs among hiv-seropositive patients with suspected tuberculosis in uganda. *Respirology* 2011;16:836-841.
13. Dawson R, Masuka P, Edwards DJ, Bateman ED, Bekker LG, Wood R, Lawn SD. Chest radiograph reading and recording system: Evaluation for tuberculosis screening in patients with advanced hiv. *Int J Tuberc Lung Dis* 2010;14:52-58.
14. van Cleeff MR, Kivihya-Ndugga LE, Meme H, Odhiambo JA, Klatser PR. The role and performance of chest x-ray for the diagnosis of tuberculosis: A cost-effectiveness analysis in nairobi, kenya. *BMC Infect Dis* 2005;5:111.
15. Shah NS, Anh MH, Thuy TT, Duong Thom BS, Linh T, Nghia DT, Sy DN, Duong BD, Chau LT, Wells C, Laserson K, Varma JK. Population-based chest x-ray screening for

pulmonary tuberculosis in people living with hiv/aids, an giang, vietnam. *Int J Tuberc Lung Dis* 2008;12:404-410.

16. Harries AD, Hargreaves NJ, Kwanjana JH, Salaniponi FM. Clinical diagnosis of smear-negative pulmonary tuberculosis: An audit of diagnostic practice in hospitals in malawi. *Int J Tuberc Lung Dis* 2001;5:1143-1147.

17. Theron G, Pooran A, Peter J, van Zyl-Smit R, Kumar Mishra H, Meldau R, Calligaro G, Allwood B, Sharma SK, Dawson R, Dheda K. Do adjunct tuberculosis tests, when combined with xpert mtb/rif, improve accuracy and the cost of diagnosis in a resource-poor setting? *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* ;40:161-168.

18. Churchyard GJ, Fielding KL, Lewis JJ, Chihota VN, Hanifa Y, Grant AD. Symptom and chest radiographic screening for infectious tuberculosis prior to starting isoniazid preventive therapy: Yield and proportion missed at screening. *AIDS* 2010;24 Suppl 5:S19-27.

19. Ling DI, Pai M, Davids V, Brunet L, Lenders L, Meldau R, Calligaro G, Allwood B, van Zyl-Smit R, Peter J, Bateman E, Dawson R, Dheda K. Are interferon- $\gamma$  release assays useful for diagnosing active tuberculosis in a high-burden setting? *Eur Respir J* 2011;38:649-656.

20. Nyboe J. Evaluation of efficiency in interpretation of chest x-ray films. *Bulletin of the World Health Organization* 1966;35:535-545.

21. Garland LH. Studies on the accuracy of diagnostic procedures. *Am J Roentgenol Radium Ther Nucl Med* 1959;82:25-38.

22. van Ginneken B, Schaefer-Prokop CM, Prokop M. Computer-aided diagnosis: How to move from the laboratory to the clinic. *Radiology* 2011;261:719-732.

23. Whitelaw A, Peter J, Sohn H, Viljoen D, Theron G, Badri M, Davids V, Pai M, Dheda K. Comparative cost and performance of light-emitting diode microscopy in hiv-tuberculosis-co-infected patients. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* 2011;38:1393-1397.
24. Shenai S, Minion J, Vadwai V, Tipnis T, Shetty S, Salvi A, Udwadia Z, Pai M, Rodrigues C. Evaluation of light emitting diode-based fluorescence microscopy for the detection of mycobacteria in a tuberculosis-endemic region. *Int J Tuberc Lung Dis* 2011;15:483-488.
25. Minion J, Pai M, Ramsay A, Menzies D, Greenaway C. Comparison of led and conventional fluorescence microscopy for detection of acid fast bacilli in a low-incidence setting. *PLoS One* 2011;6:e22495.
26. Cuevas LE, Al-Sonboli N, Lawson L, Yassin MA, Arbide I, Al-Aghbari N, Sherchand JB, Al-Absi A, Emenyonu EN, Merid Y, Okobi MI, Onuoha JO, Aschalew M, Aseffa A, Harper G, de Cuevas RM, Theobald SJ, Nathanson CM, Joly J, Faragher B, Squire SB, Ramsay A. Led fluorescence microscopy for the diagnosis of pulmonary tuberculosis: A multi-country cross-sectional evaluation. *PLoS Med* 2011;8:e1001057.
27. Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, Waugh N, Drobniewski F, Lalvani A. A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technol Assess* 2007;11:1-196.
28. Greco S, Girardi E, Navarra A, Saltini C. Current evidence on diagnostic accuracy of commercially based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis. *Thorax* 2006;61:783-790.
29. Ling DI, Flores LL, Riley LW, Pai M. Commercial nucleic-acid amplification tests for diagnosis of pulmonary tuberculosis in respiratory specimens: Meta-analysis and meta-regression. *PLoS One* 2008;3:e1536.

30. Boehme CC, Nabeta P, Henostroza G, Raqib R, Rahim Z, Gerhardt M, Sanga E, Hoelscher M, Notomi T, Hase T, Perkins MD. Operational feasibility of using loop-mediated isothermal amplification for diagnosis of pulmonary tuberculosis in microscopy centers of developing countries. *J Clin Microbiol* 2007;45:1936-1940.
31. Steingart KR, Flores LL, Dendukuri N, Schiller I, Laal S, Ramsay A, Hopewell PC, Pai M. Commercial serological tests for the diagnosis of active pulmonary and extrapulmonary tuberculosis: An updated systematic review and meta-analysis. *PLoS Med* 2011;8:e1001062.
32. Flores LL, Steingart KR, Dendukuri N, Schiller I, Minion J, Pai M, Ramsay A, Henry M, Laal S. Systematic review and meta-analysis of antigen detection tests for the diagnosis of tuberculosis. *Clinical and vaccine immunology : CVI* 2011;18:1616-1627.
33. Chang K, Lu W, Wang J, Zhang K, Jia S, Li F, Deng S, Chen M. Rapid and effective diagnosis of tuberculosis and rifampicin resistance with xpert mtb/rif assay: A meta-analysis. *The Journal of infection* 2012;64:580-588.
34. Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, Murray S, Dawson R, Whitelaw A, Hoelscher M, Sharma S, Pai M, Warren R, Dheda K. Evaluation of the xpert mtb/rif assay for the diagnosis of pulmonary tuberculosis in a high hiv prevalence setting. *Am J Respir Crit Care Med* 2011;184:132-140.
35. Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, Gler MT, Blakemore R, Worodria W, Gray C, Huang L, Caceres T, Mehdiyev R, Raymond L, Whitelaw A, Sagadevan K, Alexander H, Albert H, Cobelens F, Cox H, Alland D, Perkins MD. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the xpert mtb/rif test for diagnosis of tuberculosis and multidrug resistance: A multicentre implementation study. *Lancet* 2011;377:1495-1505.

36. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, Milovic A, Jones M, O'Brien SM, Persing DH, Ruesch-Gerdes S, Gotuzzo E, Rodrigues C, Alland D, Perkins MD. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med* 2010;363:1005-1015.
37. Leung E, Minion J, Benedetti A, Pai M, Menzies D. Microcolony culture techniques for tuberculosis diagnosis: A systematic review. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2012;16:16-23, i-iii.
38. Minion J, Pai M. Bacteriophage assays for rifampicin resistance detection in mycobacterium tuberculosis: Updated meta-analysis. *Int J Tuberc Lung Dis* 2010;14:941-951.
39. Kalantri S, Pai M, Pascopella L, Riley L, Reingold A. Bacteriophage- based tests for the detection of mycobacterium tuberculosis in clinical specimens: A systematic review and meta- analysis. *BMC Infect Dis* 2005;5:59.
40. Ling DI, Zwerling AA, Pai M. Genotype mtbdr assays for the diagnosis of multidrug-resistant tuberculosis: A meta-analysis. *Eur Respir J* 2008;32:1165-1174.
41. Morgan M, Kalantri S, Flores L, Pai M. A commercial line probe assay for the rapid detection of rifampicin resistance in mycobacterium tuberculosis: A systematic review and meta-analysis. *BMC Infect Dis* 2005;5:62.
42. Nicol MP, Zar HJ. New specimens and laboratory diagnostics for childhood pulmonary tb: Progress and prospects. *Paediatr Respir Rev* 2011;12:16-21.
43. Lawn SD, Brooks SV, Kranzer K, Nicol MP, Whitelaw A, Vogt M, Bekker LG, Wood R. Screening for hiv-associated tuberculosis and rifampicin resistance before antiretroviral therapy using the xpert mtb/rif assay: A prospective study. *PLoS Med* 2011;8:e1001067.

44. Boeree MJ, Harries AD, Godschalk P, Demast Q, Upindi B, Mwale A, Nyirenda TE, Banerjee A, Salaniponi FM. Gender differences in relation to sputum submission and smear-positive pulmonary tuberculosis in malawi. *Int J Tuberc Lung Dis* 2000;4:882-884.
45. Khan MS, Dar O, Sismanidis C, Shah K, Godfrey-Faussett P. Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: A pragmatic randomised controlled trial. *Lancet* 2007;369:1955-1960.
46. Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: A prospective study. *Lancet* 2005;365:130-134.
47. Geldenhuys HD, Kleynhans W, Buckerfield N, Tameris M, Gonzalez Y, Mahomed H, Hussey G, Hanekom W, Hatherill M. Safety and tolerability of sputum induction in adolescents and adults with suspected pulmonary tuberculosis. *Eur J Clin Microbiol Infect Dis* 2011.
48. Bell DJ, Dacombe R, Graham SM, Hicks A, Cohen D, Chikaonda T, French N, Molyneux ME, Zijlstra EE, Squire SB, Gordon SB. Simple measures are as effective as invasive techniques in the diagnosis of pulmonary tuberculosis in malawi. *Int J Tuberc Lung Dis* 2009;13:99-104.
49. Alisjahbana B, van Crevel R, Danusantoso H, Gartinah T, Soemantri ES, Nelwan RH, van der Meer JW. Better patient instruction for sputum sampling can improve microscopic tuberculosis diagnosis. *Int J Tuberc Lung Dis* 2005;9:814-817.
50. Souza Pinto V, Bammann RH. Chest physiotherapy for collecting sputum samples from hiv-positive patients suspected of having tuberculosis. *Int J Tuberc Lung Dis* 2007;11:1302-1307.

51. Morse M, Kessler J, Albrecht S, Kim R, Thakur R, Nthobatsang R, Radisowa K, Maunatlala C, Yang W, Macgregor RR, Friedman H. Induced sputum improves the diagnosis of pulmonary tuberculosis in hospitalized patients in gaborone, botswana. *Int J Tuberc Lung Dis* 2008;12:1279-1285.
52. Iriso R, Mudido PM, Karamagi C, Whalen C. The diagnosis of childhood tuberculosis in an hiv-endemic setting and the use of induced sputum. *Int J Tuberc Lung Dis* 2005;9:716-726.
53. Atiq-ur-Rehman M, Naseem A, Hussain T. Comparison of diagnostic yield of afb with sputum induction to spontaneous sputum examination in suspected pulmonary tuberculosis. *J Coll Physicians Surg Pak* 2009;19:506-509.
54. Kingkaew N, Sangtong B, Amnuaiphon W, Jongpaibulpatana J, Mankatittham W, Akksilp S, Sirinak C, Nateniyom S, Burapat C, Kittikraisak W, Monkongdee P, Varma JK. Hiv-associated extrapulmonary tuberculosis in thailand: Epidemiology and risk factors for death. *Int J Infect Dis* 2009;13:722-729.
55. Shata AM, Coulter JB, Parry CM, Ching'ani G, Broadhead RL, Hart CA. Sputum induction for the diagnosis of tuberculosis. *Arch Dis Child* 1996;74:535-537.
56. Fujita A, Murata K, Takamori M. Novel method for sputum induction using the lung flute in patients with suspected pulmonary tuberculosis. *Respirology* 2009;14:899-902.
57. Dickson SJ, Brent A, Davidson RN, Wall R. Comparison of bronchoscopy and gastric washings in the investigation of smear-negative pulmonary tuberculosis. *Clin Infect Dis* 2003;37:1649-1653.
58. Brown M, Varia H, Bassett P, Davidson RN, Wall R, Pasvol G. Prospective study of sputum induction, gastric washing, and bronchoalveolar lavage for the diagnosis of

- pulmonary tuberculosis in patients who are unable to expectorate. *Clin Infect Dis* 2007;44:1415-1420.
59. Al-Aghbari N, Al-Sonboli N, Yassin MA, Coulter JB, Atef Z, Al-Eryani A, Cuevas LE. Multiple sampling in one day to optimize smear microscopy in children with tuberculosis in yemen. *PLoS One* 2009;4:e5140.
60. Oberhelman RA, Soto-Castellares G, Gilman RH, Caviedes L, Castillo ME, Kolevic L, Del Pino T, Saito M, Salazar-Lindo E, Negron E, Montenegro S, Laguna-Torres VA, Moore DA, Evans CA. Diagnostic approaches for paediatric tuberculosis by use of different specimen types, culture methods, and pcr: A prospective case-control study. *Lancet Infect Dis* 2010;10:612-620.
61. Franchi LM, Cama RI, Gilman RH, Montenegro-James S, Sheen P. Detection of mycobacterium tuberculosis in nasopharyngeal aspirate samples in children. *Lancet* 1998;352:1681-1682.
62. Worodria W, Davis JL, Cattamanchi A, Andama A, den Boon S, Yoo SD, Hopewell PC, Huang L. Bronchoscopy is useful for diagnosing smear-negative tuberculosis in hiv-infected patients. *Eur Respir J* 2010;36:446-448.
63. Schoch OD, Rieder P, Tueller C, Altpeter E, Zellweger JP, Rieder HL, Krause M, Thurnheer R. Diagnostic yield of sputum, induced sputum, and bronchoscopy after radiologic tuberculosis screening. *Am J Respir Crit Care Med* 2007;175:80-86.
64. Willcox PA, Benatar SR, Potgieter PD. Use of the flexible fiberoptic bronchoscope in diagnosis of sputum-negative pulmonary tuberculosis. *Thorax* 1982;37:598-601.
65. Hatherill M, Hawkrigde T, Zar HJ, Whitelaw A, Tameris M, Workman L, Geiter L, Hanekom WA, Hussey G. Induced sputum or gastric lavage for community-based diagnosis of childhood pulmonary tuberculosis? *Arch Dis Child* 2009;94:195-201.

66. Nicol MP, Workman L, Isaacs W, Munro J, Black F, Eley B, Boehme CC, Zemanay W, Zar HJ. Accuracy of the xpert mtb/rif test for the diagnosis of pulmonary tuberculosis in children admitted to hospital in cape town, south africa: A descriptive study. *Lancet Infect Dis* 2011 Nov;11:819-824.
67. Hepple P, Ford N, McNerney R. Microscopy compared to culture for the diagnosis of tuberculosis in induced sputum samples: A systematic review. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2012;16:579-588.
68. Gonzalez-Angulo Y, Wiysonge CS, Geldenhuys H, Hanekom W, Mahomed H, Hussey G, Hatherill M. Sputum induction for the diagnosis of pulmonary tuberculosis: A systematic review and meta-analysis. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology* 2012;31:1619-1630.
69. Nicol MP, Zar HJ. New specimens and laboratory diagnostics for childhood pulmonary tb: Progress and prospects. *Paediatric respiratory reviews* 2011;12:16-21.
70. Lawn SD, Kerkhoff AD, Pahlana P, Vogt M, Wood R. Diagnostic yield of tuberculosis using sputum induction in hiv-positive patients before antiretroviral therapy [short communication]. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2012;16:1354-1357.
71. Parry CM, Kamoto O, Harries AD, Wirima JJ, Nyirenda CM, Nyangulu DS, Hart CA. The use of sputum induction for establishing a diagnosis in patients with suspected pulmonary tuberculosis in malawi. *Tuber Lung Dis* 1995;76:72-76.
72. Steingart KR, Ramsay A, Pai M. Optimizing sputum smear microscopy for the diagnosis of pulmonary tuberculosis. *Expert Rev Anti Infect Ther* 2007;5:327-331.

73. Perkins MD, Cunningham J. Facing the crisis: Improving the diagnosis of tuberculosis in the hiv era. *J Infect Dis* 2007;196 Suppl 1:S15-27.
74. Breslauer DN, Maamari RN, Switz NA, Lam WA, Fletcher DA. Mobile phone based clinical microscopy for global health applications. *PLoS One* 2009;4:e6320.
75. Cruciani M, Scarparo C, Malena M, Bosco O, Serpelloni G, Mengoli C. Meta-analysis of bactec mgit 960 and bactec 460 tb, with or without solid media, for detection of mycobacteria. *J Clin Microbiol* 2004;42:2321-2325.
76. WHO. Noncommercial culture and drug-susceptibility testing methods for screening patients at risk for multidrug-resistant tuberculosis. Geneva; 2010.
77. Pai M, Kalantri S, Pascopella L, Riley LW, Reingold AL. Bacteriophage-based assays for the rapid detection of rifampicin resistance in mycobacterium tuberculosis: A meta-analysis. *J Infect* 2005;51:175-187.
78. Minion J, Leung E, Menzies D, Pai M. Microscopic-observation drug susceptibility and thin layer agar assays for the detection of drug resistant tuberculosis: A systematic review and meta-analysis. *Lancet Infect Dis* 2010;10:688-698.
79. Nabeta P, Gray C, Lan Nguyen Thi N, Gotuzzo E, Dheda K, Mello F, Hillemann D, Boehme C. Evaluation of a manual nucleic acid amplification test for tuberculosis detection. IUTLD meeting. Lille, France; 2011.
80. WHO. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert mtb/rif system. Geneva, Switzerland: World Health Organisation; 2011 No. ISBN 978 92 4 150154 5.
81. Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, Kop J, Owens MR, Rodgers R, Banada P, Safi H, Blakemore R, Lan NT, Jones-Lopez EC, Levi M, Burday M, Ayakaka I, Mugerwa RD, McMillan B, Winn-Deen E, Christel L, Dailey P, Perkins MD, Persing DH,

- Alland D. Rapid detection of mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. *J Clin Microbiol* 2010;48:229-237.
82. Scott LE, McCarthy K, Gous N, Nduna M, Van Rie A, Sanne I, Venter WF, Duse A, Stevens W. Comparison of xpert mtb/rif with other nucleic acid technologies for diagnosing pulmonary tuberculosis in a high hiv prevalence setting: A prospective study. *PLoS Med* 2011;8:e1001061.
83. Rachow A, Zumla A, Heinrich N, Rojas-Ponce G, Mtafya B, Reither K, Ntinginya EN, O'Grady J, Huggett J, Dheda K, Boehme C, Perkins M, Saathoff E, Hoelscher M. Rapid and accurate detection of mycobacterium tuberculosis in sputum samples by cepheid xpert mtb/rif assay--a clinical validation study. *PLoS One* 2011;6:e20458.
84. Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, Murray S, Dawson R, Whitelaw A, Hoelscher M, Sharma S, Pai M, Warren R, Dheda K. Evaluation of the xpert mtb/rif assay for the diagnosis of pulmonary tuberculosis in a high hiv prevalence setting. *American journal of respiratory and critical care medicine* 2011;184:132-140.
85. Van Rie A, Mellet K, John MA, Scott L, Page-Shipp L, Dansey H, Victor T, Warren R. False-positive rifampicin resistance on xpert(r) mtb/rif: Case report and clinical implications. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2012;16:206-208.
86. WHO. Molecular line probe assays for rapid screening of patients at risk of multi-drug resistant tuberculosis (mdr-tb). 2008.
87. Bwanga F, Hoffner S, Haile M, Joloba ML. Direct susceptibility testing for multi drug resistant tuberculosis: A meta-analysis. *BMC Infect Dis* 2009;9:67.
88. Brossier F, Veziris N, Aubry A, Jarlier V, Sougakoff W. Detection by genotype mtbdrsl test of complex mechanisms of resistance to second-line drugs and

ethambutol in multidrug-resistant mycobacterium tuberculosis complex isolates.

*Journal of clinical microbiology* 2010;48:1683-1689.

89. Hillemann D, Rusch-Gerdes S, Richter E. Feasibility of the genotype mtbdrsl assay for fluoroquinolone, amikacin-capreomycin, and ethambutol resistance testing of mycobacterium tuberculosis strains and clinical specimens. *Journal of clinical microbiology* 2009;47:1767-1772.

90. Ignatyeva O, Kontsevaya I, Kovalyov A, Balabanova Y, Nikolayevskyy V, Toit K, Dragan A, Maxim D, Mironova S, Kummik T, Muntean I, Koshkarova E, Drobniewski F. Detection of resistance to second-line antituberculosis drugs by use of the genotype mtbdrsl assay: A multicenter evaluation and feasibility study. *Journal of clinical microbiology* 2012;50:1593-1597.

91. Grenier J, Pinto L, Nair D, Steingart K, Dowdy D, Ramsay A, Pai M. Widespread use of serological tests for tuberculosis: Data from 22 high-burden countries. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* 2012;39:502-505.

92. Dowdy DW, Steingart KR, Pai M. Serological testing versus other strategies for diagnosis of active tuberculosis in india: A cost-effectiveness analysis. *PLoS medicine* 2011;8:e1001074.

93. WHO. Commercial serodiagnostic tests for diagnosis of tuberculosis: Policy statement. Geneva: WHO; 2011.

94. Metcalfe JZ, Everett CK, Steingart KR, Cattamanchi A, Huang L, Hopewell PC, Pai M. Interferon-gamma release assays for active pulmonary tuberculosis diagnosis in adults in low- and middle-income countries: Systematic review and meta-analysis. *The Journal of infectious diseases* 2011;204 Suppl 4:S1120-1129.

95. Cattamanchi A, Smith R, Steingart KR, Metcalfe JZ, Date A, Coleman C, Marston BJ, Huang L, Hopewell PC, Pai M. Interferon-gamma release assays for the diagnosis of latent tuberculosis infection in hiv-infected individuals: A systematic review and meta-analysis. *J Acquir Immune Defic Syndr* 2011;56:230-238.
96. Sester M, Sotgiu G, Lange C, Giehl C, Girardi E, Migliori GB, Bossink A, Dheda K, Diel R, Dominguez J, Lipman M, Nemeth J, Ravn P, Winkler S, Huitric E, Sandgren A, Manissero D. Interferon-gamma release assays for the diagnosis of active tuberculosis: A systematic review and meta-analysis. *Eur Respir J* 2011;37:100-111.
97. WHO. Use of tuberculosis interferon-gamma release assays (igras) in low-and middle-income countries. Geneva; 2011.
98. Ling DI, Zwerling AA, Steingart KR, Pai M. Immune-based diagnostics for tb in children: What is the evidence? *Paediatr Respir Rev* 2011;12:9-15.
99. Mandalakas AM, Detjen AK, Hesselning AC, Benedetti A, Menzies D. Interferon-gamma release assays and childhood tuberculosis: Systematic review and meta-analysis. *Int J Tuberc Lung Dis* 2011;15:1018-1032.
100. Mortier E, Pouchot J, Girard L, Bousougant Y, Vinceneux P. Assessment of urine analysis for the diagnosis of tuberculosis. *Bmj* 1996;312:27-28.
101. Nour EM, Cherian BP, Quantrill SJ. Diagnostic yield of early morning urine samples in the diagnosis of tuberculosis. *Qjm* 2007;100:464-465.
102. Shafer RW, Kim DS, Weiss JP, Quale JM. Extrapulmonary tuberculosis in patients with human immunodeficiency virus infection. *Medicine (Baltimore)* 1991;70:384-397.
103. Monkongdee P, McCarthy KD, Cain KP, Tasaneeyapan T, Nguyen HD, Nguyen TN, Nguyen TB, Teeratakulpisarn N, Udomsantisuk N, Heilig C, Varma JK. Yield of acid-fast smear and mycobacterial culture for tuberculosis diagnosis in people with human immunodeficiency virus. *Am J Respir Crit Care Med* 2009;180:903-908.

104. Parker M. Complement fixation with urine in tuberculosis. *Am Rev Tuberc* 1931;23:733-738.
105. Chan J, Fan XD, Hunter SW, Brennan PJ, Bloom BR. Lipoarabinomannan, a possible virulence factor involved in persistence of mycobacterium tuberculosis within macrophages. *Infect Immun* 1991;59:1755-1761.
106. Briken V, Porcelli SA, Besra GS, Kremer L. Mycobacterial lipoarabinomannan and related lipoglycans: From biogenesis to modulation of the immune response. *Mol Microbiol* 2004;53:391-403.
107. Hamasur B, Bruchfeld J, Haile M, Pawlowski A, Bjorvatn B, Kallenius G, Svenson SB. Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine. *J Microbiol Methods* 2001;45:41-52.
108. Tessema TA, Hamasur B, Bjun G, Svenson S, Bjorvatn B. Diagnostic evaluation of urinary lipoarabinomannan at an ethiopian tuberculosis centre. *Scand J Infect Dis* 2001;33:279-284.
109. Boehme C, Molokova E, Minja F, Geis S, Loscher T, Maboko L, Koulchin V, Hoelscher M. Detection of mycobacterial lipoarabinomannan with an antigen-capture elisa in unprocessed urine of tanzanian patients with suspected tuberculosis. *Trans R Soc Trop Med Hyg* 2005;99:893-900.
110. Hunter SW, Gaylord H, Brennan PJ. Structure and antigenicity of the phosphorylated lipopolysaccharide antigens from the leprosy and tubercle bacilli. *J Biol Chem* 1986;261:12345-12351.
111. Daffe M, Draper P. The envelope layers of mycobacteria with reference to their pathogenicity. *Adv Microb Physiol* 1998;39:131-203.
112. Lee RE, Brennan PJ, Besra GS. Mycobacterium tuberculosis cell envelope. *Curr Top Microbiol Immunol* 1996;215:1-27.

113. Moreira LO, Mattos-Guaraldi AL, Andrade AF. Novel lipoarabinomannan-like lipoglycan (cdilam) contributes to the adherence of corynebacterium diphtheriae to epithelial cells. *Arch Microbiol* 2008;190:521-530.
114. Dheda K, Davids V, Lenders L, Roberts T, Meldau R, Ling D, Brunet L, van Zyl Smit R, Peter J, Green C, Badri M, Sechi L, Sharma S, Hoelscher M, Dawson R, Whitelaw A, Blackburn J, Pai M, Zumla A. Clinical utility of a commercial lam-elisa assay for tb diagnosis in hiv-infected patients using urine and sputum samples. *PLoS One* 2010;5:e9848.
115. Daley P, Michael JS, Hmar P, Latha A, Chordia P, Mathai D, John KR, Pai M. Blinded evaluation of commercial urinary lipoarabinomannan for active tuberculosis: A pilot study. *Int J Tuberc Lung Dis* 2009;13:989-995.
116. Mutetwa R, Boehme C, Dimairo M, Bandason T, Munyati SS, Mangwanya D, Mungofa S, Butterworth AE, Mason PR, Corbett EL. Diagnostic accuracy of commercial urinary lipoarabinomannan detection in african tuberculosis suspects and patients. *Int J Tuberc Lung Dis* 2009;13:1253-1259.
117. Reither K, Saathoff E, Jung J, Minja LT, Kroidl I, Saad E, Huggett JF, Ntinginya EN, Maganga L, Maboko L, Hoelscher M. Low sensitivity of a urine lam-elisa in the diagnosis of pulmonary tuberculosis. *BMC Infect Dis* 2009;9:141.
118. Gounder CR, Kufa T, Wada NI, Mngomezulu V, Charalambous S, Hanifa Y, Fielding K, Grant A, Dorman S, Chaisson RE, Churchyard GJ. Diagnostic accuracy of a urine lipoarabinomannan enzyme-linked immunosorbent assay for screening ambulatory hiv-infected persons for tuberculosis. *J Acquir Immune Defic Syndr* 2011;58:219-223.
119. Shah M, Variava E, Holmes CB, Coppin A, Golub JE, McCallum J, Wong M, Luke B, Martin DJ, Chaisson RE, Dorman SE, Martinson NA. Diagnostic accuracy of a urine

- lipoarabinomannan test for tuberculosis in hospitalized patients in a high hiv prevalence setting. *Journal of acquired immune deficiency syndromes* 2009;52:145-151.
120. Lawn SD, Edwards DJ, Kranzer K, Vogt M, Bekker LG, Wood R. Urine lipoarabinomannan assay for tuberculosis screening before antiretroviral therapy diagnostic yield and association with immune reconstitution disease. *AIDS* 2009;23:1875-1880.
121. Talbot E, Munseri P, Teixeira P, Matee M, Bakari M, Lahey T, von Reyn F. Test characteristics of urinary lipoarabinomannan and predictors of mortality among hospitalized hiv-infected tuberculosis suspects in tanzania. *PLoS One* 2012;7:e32876.
122. Lawn SD, Butera ST, Shinnick TM. Tuberculosis unleashed: The impact of human immunodeficiency virus infection on the host granulomatous response to mycobacterium tuberculosis. *Microbes Infect* 2002;4:635-646.
123. Doublier S, Zennaro C, Spatola T, Lupia E, Bottelli A, Deregibus MC, Carraro M, Conaldi PG, Camussi G. Hiv-1 tat reduces nephrin in human podocytes: A potential mechanism for enhanced glomerular permeability in hiv-associated nephropathy. *AIDS (London, England)* 2007;21:423-432.
124. Wood R, Racow K, Bekker LG, Middelkoop K, Vogt M, Kreiswirth BN, Lawn SD. Lipoarabinomannan in urine during tuberculosis treatment: Association with host and pathogen factors and mycobacteriuria. *BMC infectious diseases* 2012;12:47.
125. Wood R, Lawn SD. Challenges facing lipoarabinomannan urine antigen tests for diagnosing hiv-associated tuberculosis. *Expert review of molecular diagnostics* 2012;12:549-551.
126. Lawn SD, Edwards DJ, Kranzer K, Vogt M, Bekker L-G, Wood R. Urine lipoarabinomannan assay for tuberculosis screening before antiretroviral therapy diagnostic yield and association with immune reconstitution disease. *AIDS* 2009;23.

127. Badri M, Wilson D, Wood R. Effect of highly active antiretroviral therapy on incidence of tuberculosis in south africa: A cohort study. *Lancet* 2002;359:2059-2064.
128. Lawn SD, Kranzer K, Wood R. Antiretroviral therapy for control of the hiv-associated tuberculosis epidemic in resource-limited settings. *Clin Chest Med* 2009;30:685-699, viii.
129. Dheda K, Booth H, Huggett JF, Johnson MA, Zumla A, Rook GA. Lung remodeling in pulmonary tuberculosis. *J Infect Dis* 2005;192:1201-1209.
130. Lawn SD, Wilkinson RJ, Lipman MC, Wood R. Immune reconstitution and "unmasking" of tuberculosis during antiretroviral therapy. *Am J Respir Crit Care Med* 2008;177:680-685.
131. Dhasmana DJ, Dheda K, Ravn P, Wilkinson RJ, Meintjes G. Immune reconstitution inflammatory syndrome in hiv-infected patients receiving antiretroviral therapy : Pathogenesis, clinical manifestations and management. *Drugs* 2008;68:191-208.
132. Lawn SD, Myer L, Bekker LG, Wood R. Burden of tuberculosis in an antiretroviral treatment programme in sub-saharan africa: Impact on treatment outcomes and implications for tuberculosis control. *Aids* 2006;20:1605-1612.
133. Moore D, Liechty C, Ekwaru P, Were W, Mwima G, Solberg P, Rutherford G, Mermin J. Prevalence, incidence and mortality associated with tuberculosis in hiv-infected patients initiating antiretroviral therapy in rural uganda. *AIDS* 2007;21:713-719.
134. Sada E, Brennan PJ, Herrera T, Torres M. Evaluation of lipoarabinomannan for the serological diagnosis of tuberculosis. *J Clin Microbiol* 1990;28:2587-2590.
135. Sada E, Aguilar D, Torres M, Herrera T. Detection of lipoarabinomannan as a diagnostic test for tuberculosis. *J Clin Microbiol* 1992;30:2415-2418.

136. Pereira Arias-Bouda LM, Nguyen LN, Ho LM, Kuijper S, Jansen HM, Kolk AH. Development of antigen detection assay for diagnosis of tuberculosis using sputum samples. *J Clin Microbiol* 2000;38:2278-2283.
137. Patel VB, Bhigjee AI, Paruk HF, Singh R, Meldau R, Connolly C, Ndung'u T, Dheda K. Utility of a novel lipoarabinomannan assay for the diagnosis of tuberculous meningitis in a resource-poor high-hiv prevalence setting. *Cerebrospinal Fluid Res* 2009;6:13.
138. Dheda K, Van-Zyl Smit RN, Sechi LA, Badri M, Meldau R, Symons G, Khalfey H, Carr I, Maredza A, Dawson R, Wainright H, Whitelaw A, Bateman ED, Zumla A. Clinical diagnostic utility of ip-10 and lam antigen levels for the diagnosis of tuberculous pleural effusions in a high burden setting. *PLoS One* 2009;4:e4689.
139. Peter JG, Cashmore TJ, Meldau R, Theron G, van Zyl-Smit R, Dheda K. Diagnostic accuracy of induced sputum lam elisa for tuberculosis diagnosis in sputum-scarce patients. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* 2012;16:1108-1112.
140. Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, Dheda K. Urine for the diagnosis of tuberculosis: Current approaches, clinical applicability, and new developments. *Curr Opin Pulm Med* 2010;16:262-270.
141. Sechi LA, Pinna MP, Sanna A, Pirina P, Ginesu F, Saba F, Aceti A, Turrini F, Zanetti S, Fadda G. Detection of mycobacterium tuberculosis by pcr analysis of urine and other clinical samples from aids and non-hiv-infected patients. *Mol Cell Probes* 1997;11:281-285.
142. Green C, Huggett JF, Talbot E, Mwaba P, Reither K, Zumla AI. Rapid diagnosis of tuberculosis through the detection of mycobacterial DNA in urine by nucleic acid amplification methods. *Lancet Infect Dis* 2009;9:505-511.

143. Kafwabulula M, Ahmed K, Nagatake T, Gotoh J, Mitarai S, Oizumi K, Zumla A. Evaluation of pcr-based methods for the diagnosis of tuberculosis by identification of mycobacterial DNA in urine samples. *Int J Tuberc Lung Dis* 2002;6:732-737.
144. Melkonyan H, Feaver JW, Meyer E, Scheinker V, Shekhtman EM, Xin Z, Umansky S. Transrenal nucleic acids: From proof of principle to clinical tests. Problems and solutions. *Annals of the New York Academy of Sciences* 2008;1137:73-81.
145. Cannas A, Goletti D, Giradi E, Chiacchio T, Calvo L, Cuzzi G, Piacentini M, Melkonyan H, Umansky S, Lauria FN, Ippolito G, Tomei LD. *Mycobacterium tuberculosis* DNA detection in soluble fraction of urine from pulmonary tuberculosis patients. *International Journal of Tuberculosis and Lung Disease* 2008;12:146-151.
146. Shekhtman EM, Anne K, Melkonyan HS, Robbins DJ, Warsof SL, Umansky SR. Optimization of transrenal DNA analysis: Detection of fetal DNA in maternal urine. *Clin Chem* 2009;55:723-729.
147. Gopinath K, Singh S. Urine as an adjunct specimen for the diagnosis of active pulmonary tuberculosis. *Int J Infect Dis* 2009;13:374-379.
148. Garton NJ, Waddell SJ, Sherratt AL, Smith RJ, Senner C, Hinds J, Rajakumar K, Adgebola RA, Besra GS, Butcher PD, Barer MR. Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Medicine* 2008;5:e75.
149. Huggett JF, Green C, Zumla A. Nucleic acid detection and quantification in the developing world. *Biochemical Society Symposia* 2009;37:419-423.
150. Napolitano DR, Pollock N, Kashino SS, Rodrigues V, Jr., Campos-Neto A. Identification of mycobacterium tuberculosis ornithine carboamyltransferase in urine as a possible molecular marker of active pulmonary tuberculosis. *Clin Vaccine Immunol* 2008;15:638-643.

151. Fend R, Kolk AH, Bessant C, Buijtelts P, Klatser PR, Woodman AC. Prospects for clinical application of electronic-nose technology to early detection of mycobacterium tuberculosis in culture and sputum. *J Clin Microbiol* 2006;44:2039-2045.
152. Denkinger CM, Pai M. Point-of-care tuberculosis diagnosis: Are we there yet? *The Lancet infectious diseases* 2012;12:169-170.
153. Pai NP, Pai M. Point-of-care diagnostics for hiv and tuberculosis: Landscape, pipeline, and unmet needs. *Discov Med* 2012;13:35-45.
154. MSF. Towards lab-free tuberculosis diagnosis: A strategiv vision for r&d into poin-of-care testing in resource-poor settings. London; 2011.
155. Batz HG, Cooke GS, Reid SD. Towards lab-free tuberculosis diagnosis. Treatment action group, stop tb partnership. New York, USA: Imperial College London and Medecins Sans Frontieres; 2011. p. 1-36.
156. Dowdy DW, Chaisson RE, Maartens G, Corbett EL, Dorman SE. Impact of enhanced tuberculosis diagnosis in south africa: A mathematical model of expanded culture and drug susceptibility testing. *Proc Natl Acad Sci U S A* 2008;105:11293-11298.
157. Keeler E, Perkins MD, Small P, Hanson C, Reed S, Cunningham J, Aledort JE, Hillborne L, Rafael ME, Girosi F, Dye C. Reducing the global burden of tuberculosis: The contribution of improved diagnostics. *Nature* 2006;444 Suppl 1:49-57.
158. Chin CD, Laksanasopin T, Cheung YK, Steinmiller D, Linder V, Parsa H, Wang J, Moore H, Rouse R, Umvilighozo G, Karita E, Mwambarangwe L, Braunstein SL, van de Wijgert J, Sahabo R, Justman JE, El-Sadr W, Sia SK. Microfluidics-based diagnostics of infectious diseases in the developing world. *Nat Med* 2011;17:1015-1019.

159. Banday KM, Pasikanti KK, Chan EC, Singla R, Rao KV, Chauhan VS, Nanda RK. Use of urine volatile organic compounds to discriminate tuberculosis patients from healthy subjects. *Anal Chem* 2011;83:5526-5534.
160. Kolk A, Hoelscher M, Maboko L, Jung J, Kuijper S, Cauchi M, Bessant C, van Beers S, Dutta R, Gibson T, Reither K. Electronic-nose technology using sputum samples in diagnosis of patients with tuberculosis. *J Clin Microbiol* 2010;48:4235-4238.
161. Cobelens F, van den Hof S, Pai M, Squire BS, Ramsay A, Kimmerling ME. Which new diagnostics for tuberculosis, and when? *The Journal of infectious diseases* 2012;205:suppl 2.

University of Cape Town

## **Chapter 3.**

### **Sputum induction to aid the diagnosis of smear-negative or sputum-scarce TB in adults from a HIV-endemic setting**

Peter, J.G., Theron, G., Singh, N., Singh, A., and Dheda, K.

*European Respiratory Journal*. 2013 Mar 21. Epub ahead of print

#### **PhD context**

Given that up to a third of TB cases in high HIV prevalent settings are unable to spontaneously expectorate sputum, having sampling methods to obtain sputum that are technically simple to perform, safe and effective is integral to confirming a diagnosis of TB. Furthermore, an enhanced sputum sample or alternative pulmonary fluid acquisition can improve the detection of pauci-bacillary SNTB disease. Both traditional TB diagnostic technologies and novel tools, such as the MTB/RIF assay discussed in the preceding chapter require a sputum sample. SI has been found to be an easy-to-perform sampling method that is safe and thus potentially suitable for use in a primary care resource-limited setting. However, only small studies in adults with suspected SN- or SSTB in high HIV prevalence settings have been conducted to date, with widely varying culture-based diagnostic yields and accuracy measures. Furthermore, no direct comparative data is available to quantify the differences in SI performance between key patient sub-groups (e.g. HIV-infected versus uninfected). This prospective cohort was therefore enrolled to address these questions as discussed below.

## **Sputum induction to aid the diagnosis of smear-negative or sputum-scarce TB in adults from a HIV-endemic setting**

Authors: Jonathan G. Peter<sup>1</sup>, Grant Theron<sup>1</sup>, Nevanda Singh<sup>1</sup>, Avani Singh<sup>1</sup>, and Keertan Dheda<sup>1,2,3†</sup>

<sup>1</sup>Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, University of Cape Town, South Africa.

<sup>2</sup>Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, South Africa.

<sup>3</sup>Department of Infection, University College London Medical School, United Kingdom.

†Corresponding author: Keertan Dheda, +27214047651 (fax), +27214046029 (telephone), email: [Keertan.Dheda@uct.ac.za](mailto:Keertan.Dheda@uct.ac.za)

**Author contributions:** JP and KD designed the study. JP, GT, NS and AS generated the data. JP, GT and KD analysed the data and JP was the study statistician. All authors were involved in the manuscript preparation.

**Key words:** Sputum induction, HIV-infection, patient setting

**For ERJ twitter feed:** Sputum induction is unsuccessful in a fifth of patients, and culture yield is almost double in HIV-positive vs. negative and in- vs. out-patients. (Characters: 123)

Manuscript word count: 2952

## ABSTRACT

**Rationale:** Sputum induction(SI) can aid TB diagnosis, but adult data from HIV-endemic environments are limited, and it is unclear how performance varies depending on the clinical context(in-patient vs. out-patient), HIV status, and whether patients are smear-negative or sputum-scarce.

**Methods:** 696 adults with suspected smear-negative or sputum-scarce TB from Cape Town referred for routine SI. Liquid culture for *Mycobacterium tuberculosis* served as the reference standard.

**Results:** 82%(573/696) of patients provided a specimen  $\geq 1$ ml, 83%(231/278) of which had adequate quality. 15%(96/652) of SI specimens were culture-positive, and this yield was higher amongst in-patients versus out-patients[17%(71/408) vs. 10%(25/244),  $p=0.01$ ], HIV-infected versus uninfected patients [17%(51/294) vs. 9%(16/173),  $p=0.02$ ], but similar for CD4( $>200$  vs.  $< 200$ cells/ $\mu$ l) and patient(smear-negative versus sputum-scarce) sub-categories. Overall sensitivity(95%CI) of smear-microscopy was 49%(39-59), higher amongst in-patients versus out-patients[55%(43-67) vs. 32%(14-50),  $p=0.05$ ], but unaffected by HIV-co-infection, CD4 count, or patient type. 29%(203/696) of patients commenced anti-TB treatment and SI offered microbiological confirmation and susceptibility testing in only 47%(96/203).

**Conclusions:** Under programmatic conditions in an HIV-endemic environment although the yield of culture was ~2-fold higher amongst HIV-infected and in-patients, a fifth of all patients were unable to provide a specimen following sputum induction. Same-day microbiological diagnosis was only possible in ~50% of patients.

Word count: 200/200

## INTRODUCTION

The diagnosis of patients suspected of tuberculosis (TB) who are sputum smear-negative for acid-fast bacilli or who are unable to produce sputum (sputum-scarce) is a daily challenge for clinicians in HIV-endemic settings (1). In developing countries facing the dual epidemics of TB and HIV, the burden of smear-negative or sputum-scarce TB is large and accounts for approximately every second notified TB case (2). Failure to confirm a TB diagnosis negatively impacts both patients and TB control by: i) increasing morbidity and mortality (3-5), ii) fueling the transmission of multi-drug resistant TB (MDR-TB) by ineffectively treating undiagnosed disease (6, 7), and iii) exposing patients that are inappropriately given empiric TB treatment to unnecessary, toxic and prolonged drug therapy (8). Strategies to improve and decentralise the diagnosis of adults with suspected smear-negative and sputum-scarce TB are needed, with focus not only on improved diagnostic test efficacy but also on the optimisation of sputum specimen acquisition methods (9). The World Health Organisation (WHO) endorsement (10) and roll-out of the novel MTB/RIF assay looks set to offer rapid diagnostic yields close to that of solid culture techniques even in primary care settings (11-14), thus making the need to address the diagnostic bottleneck of sputum specimen acquisition urgent.

Sputum induction (SI), performed using the ultrasonic nebulisation of hypertonic saline, is a relatively simple and safe procedure suitable to use in resource-limited decentralised settings (15-20). Induction has been shown to offer similar TB case detection rates to more invasive techniques such as bronchoscopy as an aid to TB diagnosis (21, 22); It has shown particular utility for diagnostic sampling in children (23, 24) and for TB screening in asymptomatic patients prior to the initiation of anti-retrovirals (25). Consequently, advocacy for the roll-out and widespread use of sputum induction in HIV-endemic, resource-limited primary care settings is increasing.

However, data are limited for smear-negative or sputum-scarce adult TB suspects in HIV-endemic settings (15-18). Available studies are small with large variability in diagnostic performance measures and wide confidence intervals (15-18); studies are heterogenous due to differences which include: i) the clinical context of patients undergoing SI (e.g. hospital in-patient versus respiratory clinic out-patient) (19), ii) the preceding diagnostic work-up of patients prior to sampling (e.g. CXR versus no CXR) (19), iii) HIV prevalence (19), iv) sputum induction procedure (e.g. 3% vs. 5% hypertonic saline concentration) (20), and v) type of diagnostic testing on induced specimens (e.g. light vs. fluorescence microscopy) (15-18). These differences prevent useful meta-analysis (19, 20) and limit generalisability. Larger studies that provide a more robust evidence-base to guide national TB programme (NTP) policy are overdue. Furthermore, the absence of direct comparative performance data between HIV-infected and uninfected (and stratified by CD4 cell count), out- and in-patients, and smear-negative and sputum-scarce TB suspects from a single study is a major research gap, and we hypothesised that between group differences would be lower than could be expected from a simple comparison of pre-sampling TB prevalences.

To address these gaps and evaluate our hypothesis, we conducted a large, cross-sectional study of sputum induction to evaluate the procedural side-effects, sampling efficacy, specimen quality, culture-based diagnostic yield and diagnostic accuracy of smear microscopy stratified by clinical context (in-patient vs. out-patient), HIV status and CD4 count, and TB suspect reason for induction (sputum-scarce vs. smear-negative). In addition, we evaluated the ability of SI sputum sampling to allow for microbiological TB confirmation and susceptibility in all patients commencing anti-TB treatment.

## **METHODS**

### **Study population**

The study was conducted at the Groote Schuur Hospital respiratory clinic in Cape Town, South Africa. Routine sputum induction facilities are available to hospital in-patients (including ward and emergency room admissions) and out-patients (including specialist and general medical clinics) on doctors' request. Patients ( $\geq 16$  years old) referred for induction between the 12<sup>th</sup> February 2008 and the 30<sup>th</sup> May 2009 were eligible for inclusion in the study. Basic demographics, HIV status and reason for referral for induction were recorded by nursing staff. Only patients referred for induction with suspected smear-negative or sputum-scarce TB were included, and patients referred for any other indications e.g. possible *Pneumocystis jiroveci* infection or malignancy were excluded. The study was approved by the University of Cape Town Human Research Ethics Committee.

### **Diagnostic work-up and treatment**

As per routine practise, all patients referred for induction with suspected smear-negative or sputum-scarce TB had received a doctor's assessment and chest x-ray, as well as an attempted collection of a self-expectorated early morning sputum sample prior to referral. Patients with 2 x smear-negative sputum samples within 4-weeks of referral for induction were considered smear-negative. Data on the exact timing between attempted self expectoration and induction was not documented for all patients, however for sputum-scarce in-patients self-expectoration was attempted on admission. Using the laboratory and hospital pharmacy record systems, the commencement of anti-TB treatment for study patients within a month of enrolment was noted. Any patient commencing treatment without a positive TB culture result on a recent ( $\pm 4$  weeks

from enrolment) specimen (induced sputum or other) was considered to have received empiric treatment.

### **Sputum induction procedure**

Sputum induction was performed by respiratory clinic nursing staff in an enclosed negative-pressure induction booth as previously described (26). Briefly, approximately 20 ml of sterile 5% hypertonic saline (Sabax, Adcock Ingram, South Africa) was delivered via a Wilson's 402A ultrasonic nebuliser (Medimark, South Africa) over 15-20 min or until 2-4 ml of induced sputum could be collected. No pro-expectorating maneuvers were employed. Research nurses monitored for side-effects and induction was terminated if side effects developed.

### **Laboratory methods**

Induced sputa were processed by the National Health Laboratory Service reference laboratory. Specimens were decontaminated with N-acetyl-L-cysteine/NaOH and then centrifuged. Thereafter, an Auramine-O stained smear underwent fluorescence microscopy and 0.5ml of the deposit was inoculated into a Mycobacterial Growth Indicator Tube 960 (MGIT, Becton Dickinson Diagnostics, USA). Approximately half of samples received Gram staining prior to decontamination and sputum quality was determined using the Bartlett score (27). Culture-positive acid fast bacilli were identified as *M. tuberculosis* complex using either an in-house PCR method (28), or the Genotype MTBDRplus<sup>®</sup> assay (version 1; Hain LifeSciences, Nuhren, Germany) if drug susceptibility testing (DST) had been requested. Hain MTBDRplus<sup>®</sup> assay testing was introduced for routine DST only in the latter part of the study period (see Figure 1).

## Statistical analysis

All patients (or patients in a sub-group) referred for induction with suspected smear-negative or sputum-scarce TB are used as the denominator when calculating the culture-based TB case detection rates of SI. *M. tuberculosis* culture is used as the reference standard for evaluating the diagnostic accuracy of induced sputum smear microscopy and only includes patients with a valid culture result. Sensitivity, specificity, positive and negative predictive values are presented with 95% confidence intervals. Univariate and multivariate logistic regression analysis was used to investigate the predictors of induced sputum i) sputum sampling, ii) culture-positivity and iii) smear-positivity. Basic demographic and sputum induction characteristics, as well as diagnostic accuracy measures, of different patient groups were compared using  $\chi^2$ , Wilcoxon rank-sum test, and Kruskal Wallis as appropriate, and all statistical tests were 2 sided at  $\alpha=0.05$ . STATA IC, version 10 (Stata Corp, Texas, USA) was used for all statistical analyses.

## RESULTS

### Demographics, patient setting and indication for sputum induction

Of the patients referred for sputum induction during the 15 month study period, 696 patients had suspected smear-negative or sputum-scarce TB (Figure 1). 62% (434/696) and 74% (517/696) of patients referred for induction were in-patients and had suspected sputum-scarce TB, respectively. Table 1 shows basic patient demographics and the reason for sputum induction referral stratified by patient setting and HIV status. A greater proportion of out- compared to in-patients were sputum-scarce [79% (74-84) vs. 71% (67-74),  $p=0.02$ ]. The median (IQR) age of patients was 40 (32-53) years, with a younger median age amongst in- versus out-patients [38 (29-48) vs. 45 (34-57),  $p<0.001$ ], and HIV-infected vs. -uninfected patients [35 (30-42) vs. 46 (36-56),  $p<0.001$ ]. Amongst HIV-infected patients the median (IQR) CD4 cell count was 155 (65-269) cells/ $\mu\text{l}$ , with no difference in median CD4 cell count between in- and out-patients noted. 24% (74/308) of HIV-infected patients had missing CD4 cell count data.

### Sputum induction sampling efficacy, specimen quality and side-effects

Table 2 outlines sputum induction sampling efficacy, specimen quality and side-effects stratified by patient setting, HIV status and reason for referral. 82% (573/696) of referred patients successfully provided a sputum sample  $\geq 1\text{ml}$ . The success of sampling was the same irrespective of patients setting, HIV status and reason for referral for induction. A random selection of 49% (278/573) of sputum samples of  $\geq 1\text{ml}$  had a Bartlett score calculated to determine sputum quality; 83% (231/278) of these tested samples were found to be of adequate quality. Thus, 32% (220/696) of patients sputum induction was either unsuccessful or would produced a sample of sub-optimal quality. However, no association was found between sputum quality and smear- or culture-positivity [Culture-positive/all adequate sputum quality samples: 22% (8-36) 33/185 vs.

Culture-positive/all inadequate sputum quality samples: 18% (12-24) 7/32, p=0.6]. Overall, only 4% (26/696) of induced patients experienced any side-effects, of which the commonest was nausea and vomiting in nine patients.

### **Culture-based diagnostic yield and drug susceptibility testing**

The culture-based diagnostic yields of a single induced sputum sample for all patients and stratified by reason for induction, patient setting, HIV status and CD4 cell count is shown in figure 2. 6% (44/696) of patients are excluded from this analysis as, despite provision of an adequate sputum sample, the sample was rejected by the laboratory e.g. for leakage during transport or no liquid culture was requested. The overall TB culture yield (% n/N) of a single induced sputum was 15% (96/652). In a multivariate analysis, age (p=0.03), HIV positivity (p=0.05) and in-patient setting (p=0.03) were associated with TB culture-positivity (Table 3). Median (IQR) culture time-to-positivity for induced sputum samples was 13 (10-19) days, with no differences noted between patient groups. The Hain MTBDRplus<sup>®</sup> assay, introduced late in the study for routine use in accordance with national policy, was performed on 24% (23/96) of culture-positive induced sputum samples. 4% (1/23) of patients were diagnosed with MDR-TB and Isoniazid mono-resistance respectively.

### **Smear microscopy diagnostic accuracy**

Diagnostic accuracy measures of smear microscopy, using patients who were induced sputum culture-positive or -negative in the analysis, for all patients and stratified by clinical context, HIV status and CD4 cell count is shown in Table 3. Smear microscopy sensitivity (% 95% CI) was higher amongst in- versus out-patients [55% (43-67) vs. 32% (14-50), p=0.05], but unaffected by reason for referral, HIV status or CD4 cell count. Three smear microscopy positive patients were found to be culture-positive for a non-tuberculous mycobacteria.

### **Microbiological confirmation of clinical TB using a single induced sputum**

Of all patients referred for induction with suspected smear-negative or sputum-scarce TB 29% (203/696) commenced anti-TB treatment within one month of study enrolment. A single induced sputum provided culture confirmation in 47% (96/203) of cases. The proportion of culture confirmed TB amongst those receiving treatment was similar irrespective of patient setting, sputum induction referral reason and HIV status (data not shown).

University of Cape Town

## DISCUSSION

In HIV prevalent settings, both in primary and hospital practice, the burden and diagnosis of suspected smear-negative and sputum-scarce TB continues to challenge both clinicians and national TB programmes (NTPs). With the development and roll-out of novel TB diagnostic tests, such as the MTB/RIF assay, specimen acquisition has become an even more important diagnostic bottleneck. Given the simplicity, safety and performance data from large childhood TB studies, there is increasing advocacy for the widespread roll-out of sputum induction in primary care clinics in resource-limited, HIV-endemic settings. Yet, for adults with suspected smear-negative or sputum-scarce TB, data are limited and heterogeneous. Thus, the main findings of our study are of relevance to both clinicians and NTPs, and include: i) under programmatic conditions, sputum induction was unsuccessful or samples were of sub-optimal quality 32% of cases, ii) the overall culture-based diagnostic yield of a single induced sputum was 16%, and culture-yield was almost 2-fold higher amongst HIV-infected and in-patients, iii) the overall sensitivity of induced sputum smear microscopy is only 49%, but is higher amongst in- compared to out-patients, and iv) irrespective of clinical setting, HIV status or patient phenotype, a single induced sputum sample only microbiologically-confirms TB offering susceptibility testing in just under half of all patients receiving anti-TB treatment,

Small studies conducted under strict research conditions over-estimate procedural success rates and diagnostic accuracy and are prone to reporting bias (29). Thus, data from this large programmatic cohort has important implications for clinicians and TB services managing smear-negative and sputum-scarce TB suspects in HIV-endemic settings. Notably, the main outcome measures of diagnostic utility, including procedural success rate, culture-based yield, smear microscopy sensitivity, and the ability to offer a microbiologically confirmed TB diagnosis for drug susceptibility testing, were all substantially lower, or on the lower-end compared to those

previously reported (15-19). Furthermore, we have unpublished data that suggests that the sensitivity of the Xpert MTB/RIF assay is reduced in induced sputum specimens from adults with smear-negative and sputum-scarce TB, which is plausible given the paucibacillary nature of TB disease in these patients. These programmatic findings also have relevance to low-burden TB settings where MTB/RIF testing is increasingly routine and sputum sampling is thus the major diagnostic challenge. This highlights the limitations of sputum induction and emphasises the need for ongoing research to improve sampling methodology and to investigate and utilise non-sputum based biological fluids in the diagnosis of adult smear-negative and sputum-scarce TB in both HIV-endemic and low TB burden settings.

How does the performance of smear and culture differ on induced versus self-expectorated sputum? Although this study did not directly compare performance in the same hospital setting, data from large recent out-patient studies of TB suspects presenting to out-patient settings in Cape Town, find culture-based diagnostic yields between 24-28% (11, 12), approximately 3-fold higher than the induced sputum samples from out-patients this study. Smear microscopy sensitivity in induced sputum specimens is also, unsurprisingly, reduced compared to in self-expectorated sputum specimens, although in HIV-infected patients sensitivity appears equivalent at ~50% (11). This highlights the particular utility of sputum induction for HIV-infected patients with suspected TB irrespective of self-expectorated smear status. In fact, the incremental culture-based yield of induced sputum, even in HIV-infected patients able to self-expectorate, has been recently demonstrated (30).

Does induction offer greater diagnostic utility in certain patient sub-groups, and is the only important determinant of these differences related to the expected differences in pre-sampling TB prevalence between groups? Our study finding of an almost 2-fold increased culture-yield in HIV-infected compared to uninfected and in- compared to out-patients confirms that induction

offers particular utility in these two patient sub-groups. However, as hypothesised, the between group differences found were lower than expected given that studies from a similar setting have shown a 4-5 fold higher TB prevalence amongst HIV-infected compared to un-infected patients (31). The lower than expected differences in culture-yield are not explained by procedure-related factors such as sampling success and side-effects which we found to be unaffected by HIV-status or clinical context. Thus, these differences likely reflect simply the limitations of using a single sputum sample for TB diagnosis in HIV-infected patients with high rates of disseminated, extrapulmonary TB and pauci-bacillary pulmonary disease. In addition, it is worth noting that despite the increased culture-yield in the above patient groups, we found that in all patient groups sputum induction could only microbiologically confirm TB allowing for drug susceptibility testing in under 50% of patients commencing anti-TB treatment. Undoubtedly, further large prospective studies in HIV-endemic settings with high rates of empiric treatment are required.

Increased TB case detection at primary care level offers a number of potential public health benefits including: i) reduction in diagnostic delay and consequently TB-related transmission, morbidity and even mortality benefits (1, 32), ii) early and increased diagnosis of MDR-TB, and iii) reduction in empiric treatment rates and hence inappropriate toxic drug exposure. However, should sputum induction facilities be decentralised and rolled-out to primary care facilities to aid the diagnosis of adult smear-negative and sputum-scarce TB? Unfortunately, no impact data evaluating the effect of sampling with sputum induction on patient important outcomes such as time-to-treatment and rates of treatment initiation are available for any patient group. However, data from large cohorts of children with suspected TB and from anti-retroviral initiation clinics support the use of induction (23-25). However, for adults presenting with suspected smear-negative or sputum-scarce TB to primary care clinics few primary care data are available.

Outcome data from the out-patient group in our study raise concern that the benefits of induction in primary clinics may be limited as, not only was the culture-based detection less than 10%, but

the sensitivity of smear microscopy was significantly lower in out-compared to in-patients. Furthermore, a study from Malawi found that a healthcare worker observed self-expectorated sputum sample offered the greatest initial diagnostic yield with minimal additional benefit from the use of sputum induction (15). Pragmatic studies evaluating the impact of sputum induction on patient-important outcomes such as treatment initiation, studies directly comparing induction to other simple sputum sampling strategies e.g. healthcare-worker instruction, and studies evaluating the performance of the novel MTB/RIF assay on induced sputum samples, are urgently need from primary care clinic settings.

This study has important limitations. Missing data was a problem given the programmatic nature of the study. Approximately a quarter of patients had an unknown or refused HIV status, and a similar proportion of HIV-infected patients had missing recent CD4 cell counts. However, induction data (reason for referral, sampling, side-effects) of the HIV unknown group was similar to the HIV uninfected, and despite missing data, large numbers of both HIV-infected and uninfected patients were included in whom sputum induction performance could be evaluated. Limited clinical data, other than the reason for referral, and no radiological data was available for study patients, and in addition, no clinical follow-up data was available. This lack of more detailed data makes it difficult to further explore the reasons for failed induction, as well as to assess for the appropriateness of empiric treatment, treatment response or the development of MDR-TB. Lastly the lack of a direct comparator method is a limitation and we are thus only able to compare our study findings with yield of self-expectorated sputum from other studies performed in a similar setting.

In conclusion, this study provides robust outcome data on the routine diagnostic utility of sputum induction for adults with suspected smear-negative or sputum-scarce TB. Although SI was safe and culture yield was almost 2-fold higher amongst HIV-infected or in-patients, in almost a third

of cases SI was unsuccessful or samples of sub-optimal quality were obtained. Same-day diagnosis using smear microscopy was only possible in less than half of patients. Direct comparison between patient sub-groups, with the low culture-yield and same-day smear microscopy diagnosis in out-patients undergoing sputum induction, raises concern about the use of sputum induction as the preferred initial sputum sampling strategy for adult out-patients with suspected smear-negative or sputum-scarce TB. Impact studies, and studies comparing different initial sputum sampling strategies to aid diagnosis in adults with suspected smear-negative and sputum-scarce TB presenting to primary care facilities in HIV-endemic settings are urgently required to inform policy.

University of Cape Town

## **ACKNOWLEDGEMENTS**

We thank the respiratory clinic nursing staff who were involved in the recruitment of patients and the performance of sputum induction. We would like to acknowledge Dr Liana Roodt and Mr Toby Tebbutt for data capture. Their support was greatly appreciated.

## **COMPETING INTERESTS AND FUNDING**

JP was supported by the Fogarty International Clinical Research Scholars/Fellows Support Centre National Institutes of Health grant R24TW007988, SATBAT the Discovery Foundation and the EDCTP for this project. GT is supported by the EDCTP, the NRF and the Claude Leon Foundation. KD is supported by the EDCTP (TB NEAT and TESA) and the SA DST and NRF (SARChI). None of the authors have any conflicts of interest to declare.

## REFERENCES

1. Whitehorn J, Ayles H, Godfrey-Faussett P. Extra-pulmonary and smear-negative forms of tuberculosis are associated with treatment delay and hospitalisation. *Int J Tuberc Lung Dis.* 2010 Jun;14(6):741-4.
2. WHO. WHO Report 2011: Global tuberculosis control. Geneva: World Health Organisation 2011. Report No.: ISBN 978 92 4 156438 0.
3. Banda H, Kang'ombe C, Harries AD, Nyangulu DS, Whitty CJ, Wirima JJ, et al. Mortality rates and recurrent rates of tuberculosis in patients with smear-negative pulmonary tuberculosis and tuberculous pleural effusion who have completed treatment. *Int J Tuberc Lung Dis.* 2000 Oct;4(10):968-74.
4. Harries AD, Nyirenda TE, Banerjee A, Boeree MJ, Salaniponi FM. Treatment outcome of patients with smear-negative and smear-positive pulmonary tuberculosis in the National Tuberculosis Control Programme, Malawi. *Trans R Soc Trop Med Hyg.* 1999 Jul-Aug;93(4):443-6.
5. Salaniponi FM, Gausi F, Kwanjana JH, Harries AD. Time between sputum examination and treatment in patients with smear-negative pulmonary tuberculosis. *Int J Tuberc Lung Dis.* 2000 Jun;4(6):581-3.
6. Behr MA, Warren SA, Salamon H, Hopewell PC, Ponce de Leon A, Daley CL, et al. Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet.* 1999 Feb 6;353(9151):444-9.
7. Tostmann A, Kik SV, Kalisvaart NA, Sebek MM, Verver S, Boeree MJ, et al. Tuberculosis transmission by patients with smear-negative pulmonary tuberculosis in a large cohort in the Netherlands. *Clin Infect Dis.* 2008 Nov 1;47(9):1135-42.

8. McIlleron H, Meintjes G, Burman WJ, Maartens G. Complications of antiretroviral therapy in patients with tuberculosis: drug interactions, toxicity, and immune reconstitution inflammatory syndrome. *The Journal of infectious diseases*. 2007 Aug 15;196 Suppl 1:S63-75.
9. Getahun H, Harrington M, O'Brien R, Nunn P. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: informing urgent policy changes. *Lancet*. 2007 Jun 16;369(9578):2042-9.
10. WHO. AUTOMATED REAL-TIME NUCLEIC ACID AMPLIFICATION TECHNOLOGY FOR RAPID AND SIMULTANEOUS DETECTION OF TUBERCULOSIS AND RIFAMPICIN RESISTANCE: Xpert MTB/RIF SYSTEM. Geneva, Switzerland: World Health Organisation 2011. Report No.: ISBN 978 92 4 150154 5.
11. Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet*. 2011 Apr 30;377(9776):1495-505.
12. Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, Murray S, et al. Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. *American journal of respiratory and critical care medicine*. 2011 Jul 1;184(1):132-40.
13. Weyer K, Mirzayev F, Migliori G, Van Gemert W, D'Ambrosio L, Zignol M, et al. Rapid molecular TB diagnosis: evidence, policy-making and global implementation of Xpert(R)MTB/RIF. *Eur Respir J*. 2012 Nov 22.
14. Pantoja A, Fitzpatrick C, Vassall A, Weyer K, Floyd K. Xpert MTB/RIF for diagnosis of TB and drug-resistant TB: a cost and affordability analysis. *Eur Respir J*. 2012 Dec 20.
15. Bell DJ, Dacombe R, Graham SM, Hicks A, Cohen D, Chikaonda T, et al. Simple measures are as effective as invasive techniques in the diagnosis of pulmonary tuberculosis in Malawi. *Int J Tuberc Lung Dis*. 2009 Jan;13(1):99-104.

16. Wilson D, Nachege J, Morroni C, Chaisson R, Maartens G. Diagnosing smear-negative tuberculosis using case definitions and treatment response in HIV-infected adults. *Int J Tuberc Lung Dis.* 2006 Jan;10(1):31-8.
17. Morse M, Kessler J, Albrecht S, Kim R, Thakur R, Nthobatsang R, et al. Induced sputum improves the diagnosis of pulmonary tuberculosis in hospitalized patients in Gaborone, Botswana. *Int J Tuberc Lung Dis.* 2008 Nov;12(11):1279-85.
18. Parry CM, Kamoto O, Harries AD, Wirima JJ, Nyirenda CM, Nyangulu DS, et al. The use of sputum induction for establishing a diagnosis in patients with suspected pulmonary tuberculosis in Malawi. *Tuber Lung Dis.* 1995 Feb;76(1):72-6.
19. Hepple P, Ford N, McNerney R. Microscopy compared to culture for the diagnosis of tuberculosis in induced sputum samples: a systematic review. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease.* 2012 May;16(5):579-88.
20. Gonzalez-Angulo Y, Wiysonge CS, Geldenhuys H, Hanekom W, Mahomed H, Hussey G, et al. Sputum induction for the diagnosis of pulmonary tuberculosis: a systematic review and meta-analysis. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology.* 2012 Nov 18;31(7):1619-30.
21. Anderson C, Inhaber N, Menzies D. Comparison of sputum induction with fiber-optic bronchoscopy in the diagnosis of tuberculosis. *Am J Respir Crit Care Med.* 1995 Nov;152(5 Pt 1):1570-4.
22. Conde MB, Soares SL, Mello FC, Rezende VM, Almeida LL, Reingold AL, et al. Comparison of sputum induction with fiberoptic bronchoscopy in the diagnosis of tuberculosis: experience at an acquired immune deficiency syndrome reference center in Rio de Janeiro, Brazil. *Am J Respir Crit Care Med.* 2000 Dec;162(6):2238-40.

23. Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: a prospective study. *Lancet*. 2005 Jan 8-14;365(9454):130-4.
24. Hatherill M, Hawkrigde T, Zar HJ, Whitelaw A, Tameris M, Workman L, et al. Induced sputum or gastric lavage for community-based diagnosis of childhood pulmonary tuberculosis? *Arch Dis Child*. 2009 Mar;94(3):195-201.
25. Lawn SD, Kerkhoff AD, Pahlana P, Vogt M, Wood R. Diagnostic yield of tuberculosis using sputum induction in HIV-positive patients before antiretroviral therapy. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*. 2012 Oct;16(10):1354-7.
26. Cashmore TJ, Peter JG, van Zyl-Smit RN, Semple PL, Maredza A, Meldau R, et al. Feasibility and diagnostic utility of antigen-specific interferon-gamma responses for rapid immunodiagnosis of tuberculosis using induced sputum. *PLoS One*. 2010;5(4):e10389.
27. Bartlett RC. *Medical Microbiology: Quality, Cost and Clinical Relevance*. New York: John Wiley and Sons; 1974.
28. De Wit D, Steyn L, Shoemaker S, Sogin M. Direct detection of *Mycobacterium tuberculosis* in clinical specimens by DNA amplification. *Journal of clinical microbiology*. 1990 Nov;28(11):2437-41.
29. Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. *BMJ*. 2002 Feb 23;324(7335):477-80.
30. Lawn SD, Kerkhoff AD, Pahlana P, Vogt M, Wood R. Diagnostic yield of tuberculosis using sputum induction in HIV-positive patients before antiretroviral therapy. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*. 2012 Oct;16(10):1354-7.

31. Gupta A, Wood R, Kaplan R, Bekker LG, Lawn SD. Tuberculosis incidence rates during 8 years of follow-up of an antiretroviral treatment cohort in South Africa: comparison with rates in the community. *PLoS One*. 2012;7(3):e34156.
32. Finnie RK, Khoza LB, van den Borne B, Mabunda T, Abotchie P, Mullen PD. Factors associated with patient and health care system delay in diagnosis and treatment for TB in sub-Saharan African countries with high burdens of TB and HIV. *Trop Med Int Health*. 2011 Apr;16(4):394-411.

University of Cape Town

## FIGURE LEGEND

### Figure 1.

Study flow diagram of patients undergoing sputum induction and their smear and culture results

§202/696 patients had an unknown HIV status and 186/696 were HIV-uninfected.

¶No recent CD4 cell count results were available for 74/308 HIV infected patients

\*Only patients that had a induced sputum culture performed are included as the denominator.

ζRoutine genotypic DST using the Hain MTBDRplus assay was not starting until the last few months of the study enrollment period.

### Figure 2.

Overall diagnostic yield of TB culture and a single induced sputum stratified by reason for sputum induction, out- and in-patient setting, HIV status and CD4 cell count

†Denominator incorporated all patients undergoing sputum induction (n=696) including those in whom sputum induction failed to produce a sample for diagnostic testing (n=123), but excluding those in whom no culture was requested/sample was rejected for leakage (n=44). See figure 1 for further detail.

Figure 1.

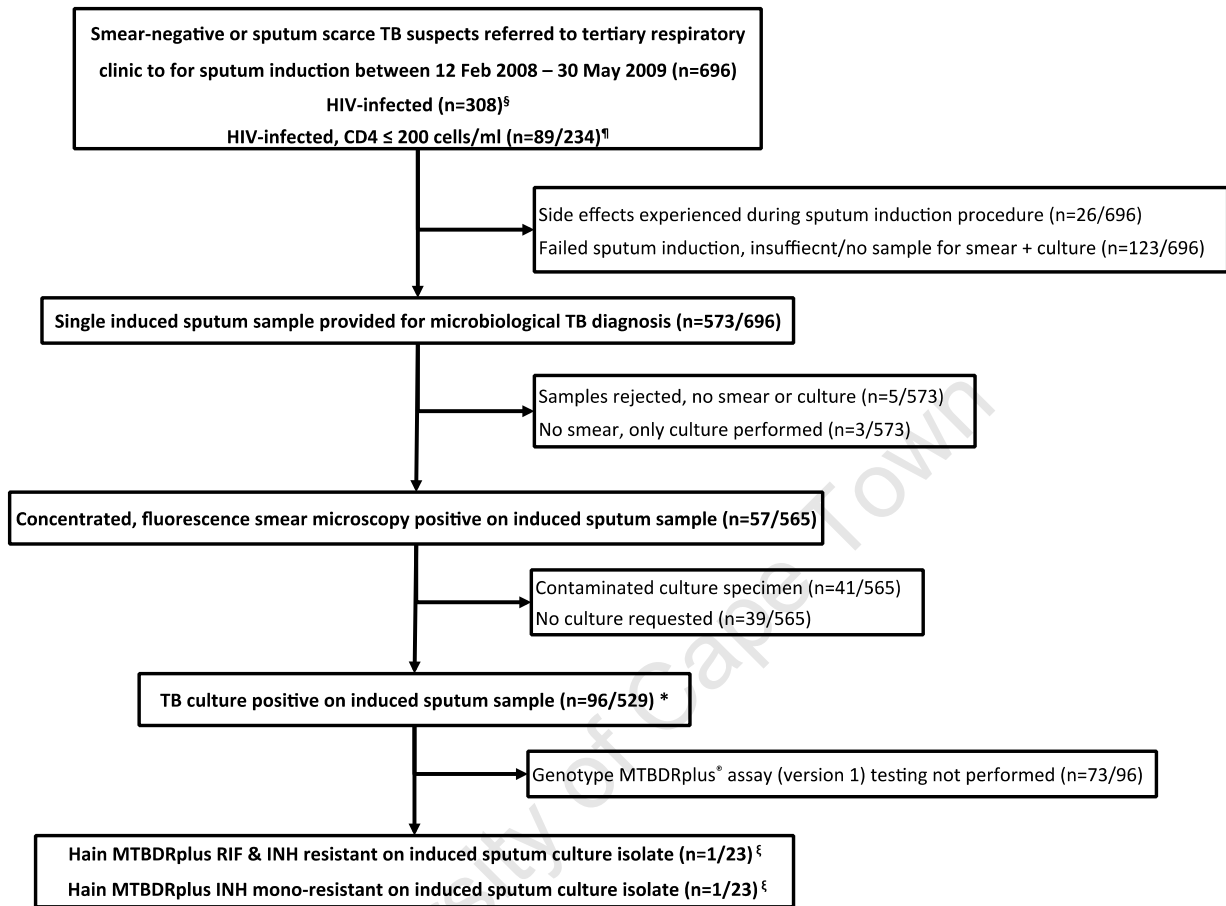
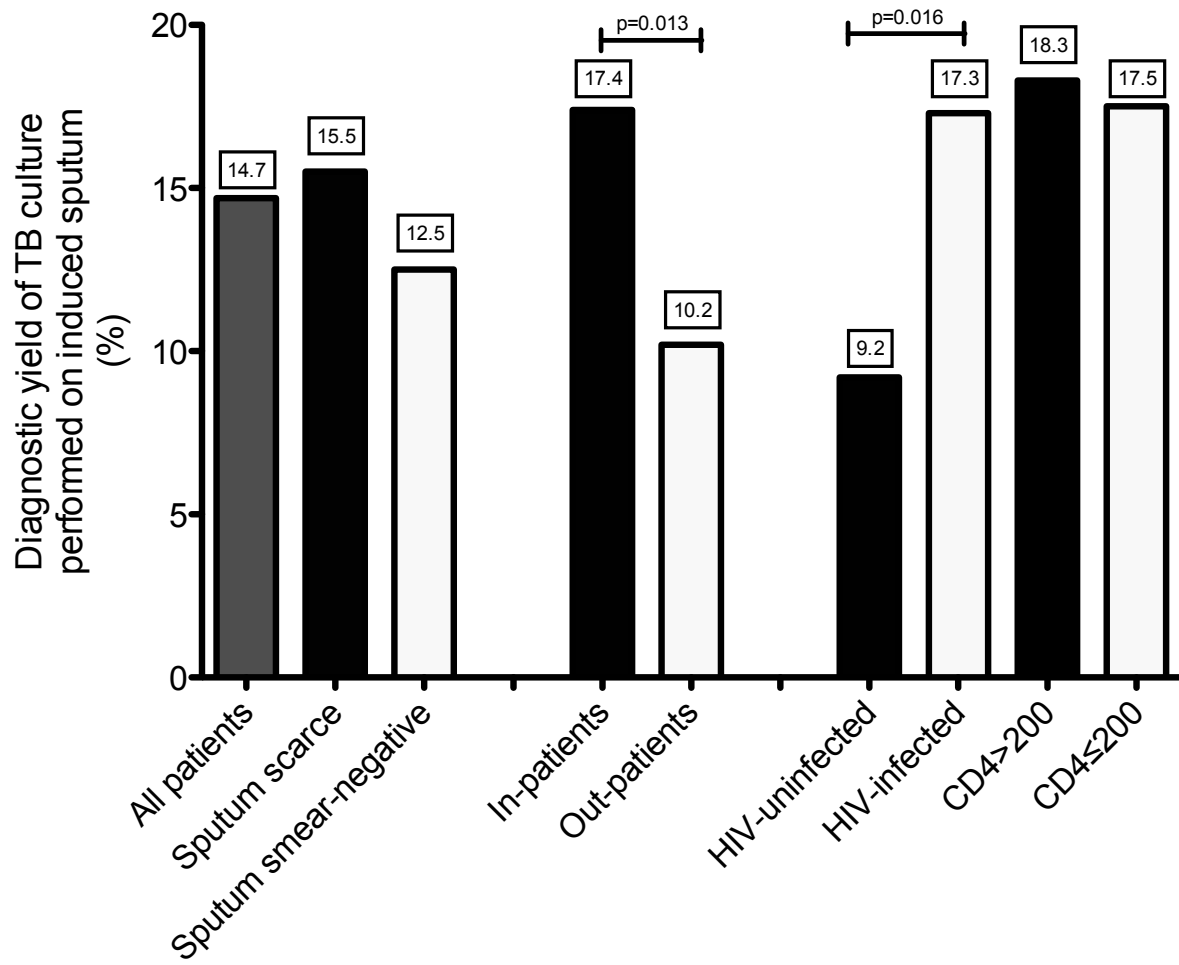


Figure 2



**Table 1.** Patients demographics and indication for referral in patient undergoing sputum induction stratified by patient setting and HIV status.

<b>Demographic and clinical characteristics</b>							
<b>Patient Characteristic</b>	<b>All (N=696)</b>	<b>Out-patient (N=262)</b>	<b>In-patient (N=434)</b>	<b>HIV-infected (N=308)</b>	<b>HIV-uninfected (N=186)</b>	<b>HIV status unknown (N=202)</b>	<b>p-value<sup>§</sup></b>
<b>Median (IQR) age, years</b>	40 (32-53)	45 (34-57)*	38 (29-48)*	35 (30-42)* <sup>1</sup>	46 (36-56)* <sup>1</sup>	49 (33-61)* <sup>1</sup>	*p<0.001 * <sup>1</sup> p<0.001
<b>Female, n(%)</b>	408 (58)	152 (58)	256 (59)	192 (62)	98 (53)	118 (58)	n/s
<b>Median (IQR) CD4 cell count<sup>†</sup>, cells/<math>\mu</math>l</b>	155 (65-269)	185 (74-290)	146 (60-262)	155 (65-269)	n/a	n/a	n/s
<b>Reason for referral to undergo sputum induction</b>							
<b>Unable to produce (sputum scarce), n(%)</b>	517 (74)	208 (79)*	309 (71)*	228 (74)	133 (72)	156 (77)	*p=0.02
<b>Smear negative x 2, n(%)</b>	179 (26)	54 (21)*	125 (29)*	80 (26)	53 (28)	46 (23)	*p=0.02

<sup>†</sup>74 HIV-infected patients missing CD4 cell count data

<sup>§</sup>P-values indicate significant differences between different patient groups (\* indicate the two or three groups compared)

n/a: not applicable; n/s: not significant (p>0.05)

**Table 2.** Sampling outcomes for sputum induction using a Wilson ultrasonic nebuliser and 5% hypertonic saline

Sampling characteristic	Patient group						
	All <sup>†</sup> % (n/N)	Out-patient % (n/N)	In-patient % (n/N)	HIV infected <sup>‡</sup> % (n/N)	HIV uninfected <sup>‡</sup> % (n/N)	Sputum scarce % (n/N)	Smear-negative x2 % (n/N)
<b>Sputum sample ≥1ml produced after induction</b>	82 573/696	86 (224/262)	80 (349/434)	79 (244/308)	86 (160/186)	81 (419/517)	86 (154/179)
<b>Side-effects experience during procedure</b>	4 (26/696)	2 (6/262)	5 (20/434)	4 (13/308)	4 (7/186)	3 (16/517)	6 (10/179)
<b>Sputum sample of adequate quality<sup>†</sup></b>	83 (231/278)	78 (87/111)	86 (144/167)	92 <sup>*1</sup> (81/88)	81 <sup>*1</sup> (80/99)	80 <sup>*2</sup> (155/194)	91 <sup>*2</sup> (76/84)

<sup>†</sup>A random selection of ~50% of the induced sputum samples received by the National Health Laboratory Services had a gram stain performed and a Bartlett score was determined (n=278/573). A Bartlett score of ≥0 is considered to be an adequate quality sputum sample for bacteriology.

<sup>‡</sup>Only patients with available HIV results are included (202 patients test refused/hiv unknown)

P-values show indicate significant differences between patient groups for a particular diagnosis accuracy measure (\* indicates the two patient groups compared); specific p-values: <sup>\*1</sup>p=0.02; <sup>\*2</sup>p=0.02

Table 3. Predictors of induced sputum TB culture positivity in a univariate and multivariate regression analysis<sup>‡</sup>.

Patient characteristic (s)	Odds ratio (95% CI)	P-value
<b>Univariate</b>		
Age (years)	0.97 (0.95-0.99)	<0.001
Female	0.88 (0.56-1.4)	0.6
HIV positive <sup>#</sup>	2.4 (1.3-4.4)	0.005
CD4 count (if HIV positive)	0.999 (0.996-1.001)	0.4
Sputum-scarce	1.3 (0.8-2.3))	0.3
In-patient	2.1 (1.3-3.5)	0.003
Adequate sputum quality	0.78 (0.31-1.94)	0.6
<b>Multivariate*</b>		
Age (years)	0.972 (0.947-0.997)	0.03
HIV positive	1.88 (0.99-3.54)	0.05
In-patient	2.06 (1.07-3.93)	0.03

<sup>‡</sup>All patients with a either a TB culture positive or negative induced sputum result were included (N=488). Patients unable to produce an induced sputum sample or with a contaminated culture result were excluded.

<sup>#</sup>Missing HIV results meant that N=348 for this univariate analysis and for the subsequent multivariate analyses

\*Only variable with a significant p-value are shown for the multivariate analyses

**Table 4.** Diagnostic accuracy of induced sputum smear microscopy stratified by patient setting, HIV status and CD4 cell count. *M. tuberculosis* culture positivity is used as the reference standard.

<b>Patient group (%, 95% CI)</b>	<b>All<sup>†</sup> N=485</b>	<b>Out-patient N=190</b>	<b>In-patient N=295</b>	<b>HIV infected<sup>‡</sup> N=210</b>	<b>HIV uninfected<sup>‡</sup> N=136</b>	<b>CD4&gt;200<sup>§</sup> N=104</b>	<b>CD4≤200<sup>§</sup> N=89</b>
<b>Sensitivity</b>	49 (39-59) 47/96	32 <sup>*1</sup> (14-50) 8/25	55 <sup>*1</sup> (43-67) 39/71	53 (40-66) 27/51	50 (28-72) 8/16	58 (39-76) 14/24	50 (31-69) 12/24
<b>Specificity</b>	98 (97-100) 383/389	99 (97-100) 164/165	98 (95-99) 219/224	98 (94-99) 155/159	100 (97-100) 120/120	99 (93- 100) 79/80	95 (87-98) 62/65
<b>PPV</b>	89 (77-95) 47/53	89 (57-98) 8/9	89 (76-95) 39/44	87 (71-95) 27/31	100 (68-100) 8/8	93 (70-99) 14/15	80 (55-93) 12/15
<b>NPV</b>	89 (85-91) 383/432	91 (86-94) 164/181	87 (83-91) 219/251	87 (81-92) 155/179	94 (88-97) 120/128	89 (81-94) 79/89	84 (74-91) 62/74

<sup>†</sup>Only patients with a valid induced sputum smear and culture result are included (see Figure 1)

<sup>‡</sup>Only patients with available HIV results are include (202 patients test refused/hiv unknown)

<sup>§</sup>Only patients with available CD4 cell count data are include (74 HIV-infected patients missing recent CD4 cell count data)

P-values show indicate significant differences between patient groups for a particular diagnosis accuracy measure (\* indicates the two patient groups compared); specific p-values: <sup>\*1</sup>p=0.05

## **Chapter 4.**

### **Comparison of two methods for acquisition of sputum samples for diagnosis of suspected tuberculosis in smear-negative or sputum-scarce people: a randomised controlled trial**

Peter, J.G., Theron, G., Pooran, A., Thomas, J., Pascoe, M. and Dheda, K.

*Lancet Respiratory Medicine*. 2013, July 19. Epub ahead of print

#### **PhD context**

As discussed in the previous chapter, SI offers utility as an aid to the diagnosis of SN- or SSTB, and given its promising performance characteristics especially in children, there is increasing advocacy for the use and widespread roll-out of SI capacity in primary care clinics in SA. However, given the findings outlined in the previous chapter, the utility of SI in outpatients is modest. The comparative benefits of SI as apposed to even simpler sampling strategies such as healthcare worker-provided instruction – particularly on patient-important treatment outcomes beyond simple culture-based diagnosis and smear microscopy sensitivity – is unclear. Therefore, we undertook a randomised controlled clinical trial to evaluate the preferred sputum sampling method for adults with suspected SN- or SSTB presenting to primary care clinics in Cape Town, South Africa.

# Comparison of two methods for acquisition of sputum samples for diagnosis of suspected tuberculosis in smear-negative or sputum-scarce people: a randomised controlled trial



Jonathan G Peter, Grant Theron, Anil Pooran, Johnson Thomas, Mellissa Pascoe, Keertan Dheda

## Summary

**Background** Sputum obtained either under instruction from a health-care worker or through induction can improve case detection of active tuberculosis. However, the best initial sputum sampling strategy for adults with suspected smear-negative or sputum-scarce tuberculosis in primary care is unclear. We compared these two methods of sample acquisition in such patients.

**Methods** In this randomised controlled trial, we enrolled adults (age  $\geq 18$  years) with sputum-scarce or smear-negative suspected tuberculosis from three primary care clinics in Cape Town, South Africa. Patients were randomly assigned (1:1) to receive either health-care worker instruction or induction to obtain sputum samples. Neither patients nor investigators were masked to allocation. The primary outcome was the proportion of patients who had started treatment after 8 weeks in a modified intention-to-treat population. Secondary outcomes were proportions starting treatment within different time periods, proportion of patients producing sputum for diagnosis, adverse effects, sputum samples' quality, and case detection by diagnostic method. This study is registered with ClinicalTrials.gov, number NCT01545661.

**Findings** We enrolled 481 patients, of whom 213 were assigned to health-care worker instruction versus 268 assigned to induction. The proportion of patients who started treatment in the 8 weeks after enrolment did not differ significantly between groups (53/213 [25%] vs 73/268 [27%]; OR 0.88, 95% CI 0.57–1.36;  $p=0.56$ ). A higher proportion of instructed versus induced patients initiated empiric treatment based on clinical and radiography findings (32/53 [60%] vs 28/73 [38%];  $p=0.015$ ). An adequate sputum sample  $\geq 1$  mL was acquired in a lower proportion of instructed versus induced patients (164/213 [77%] vs 238/268 [89%];  $p<0.0001$ ), and culture-based diagnostic yield was lower in instructed versus induced patients (24/213 [11%] vs 51/268 [19%];  $p=0.020$ ). However, same-day tuberculosis case detection was similar in both groups using either smear microscopy (13/213 [6%] vs 22/268 [8%];  $p=0.38$ ) or Xpert-MTB/RIF assay (13/89 [15%] vs 20/138 [14%];  $p=0.98$ ). No serious adverse events occurred in either group; side-effects related to sample acquisition were reported in 32 of 268 (12%) patients who had sputum induction and none who had instruction. Cost per procedure was lower for instructed than for induced patients (US\$2.14 vs US\$7.88).

**Interpretation** Although induction provides an adequate sample and a bacteriological diagnosis more frequently than instruction by a health-care worker, it is more costly, does not result in a higher proportion of same-day diagnoses, and—because of widespread empiric treatment—may not result in more patients starting treatment. Thus, health-care worker instruction might be the preferred strategy for initial collection of sputum samples in adults with suspected sputum-scarce or smear-negative tuberculosis in a high burden primary care setting.

**Funding** South African National Research Foundation, European Commission, National Institutes of Health, European and Developing Countries Clinical Trials Partnership, Discovery Foundation.

## Introduction

Tuberculosis kills more than 1 million people in Africa every year.<sup>1</sup> Several hurdles hamper effective control of tuberculosis, but an inability to access new and accurate diagnostic instruments is a major unmet need that is crucial to achieving the Millennium Development Goals.<sup>2</sup> Confirmation of tuberculosis requires not only an effective diagnostic test, but also acquisition of a biological sample of adequate volume and quality. Thus, obtaining such a sample is as important as having access to an accurate diagnostic device, especially in regions with high HIV prevalence, where most notified cases of tuberculosis are smear-negative or sputum-

scarce (patient is unable to produce sputum),<sup>1</sup> use of empiric tuberculosis treatment is common, and tuberculosis-related mortality is high.<sup>3</sup> Moreover, people with a negative smear are more likely to be admitted to hospital and have delays in diagnosis than are people with a positive smear.<sup>3</sup> WHO has recommended the Xpert MTB/RIF assay (Cepheid; CA, USA) for the frontline diagnosis of active tuberculosis in people with HIV.<sup>4</sup> However, around one in five people with both HIV and tuberculosis will have a negative result,<sup>5–7</sup> and this assay can only be used in patients who are able to provide a sputum sample. Thus, despite the advent of new diagnostic techniques, interventions to improve sample

Published Online

July 19, 2013

[http://dx.doi.org/10.1016/S2213-2600\(13\)70120-6](http://dx.doi.org/10.1016/S2213-2600(13)70120-6)

See Online/Comment

[http://dx.doi.org/10.1016/S2213-2600\(13\)70144-9](http://dx.doi.org/10.1016/S2213-2600(13)70144-9)

See Online for an interview with Jonathan Peter

Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, Faculty of Health Sciences (J G Peter MBChB, G Theron PhD, A Pooran PhD, J Thomas MBChB, M Pascoe MBChB, Prof K Dheda MBChB PhD), Institute of Infectious Diseases and Molecular Medicine (Prof K Dheda), University of Cape Town, Cape Town, South Africa; TB Vaccine Group, Jenner Institute, University of Oxford, Oxford, UK (J G Peter); and Department of Infection, University College London Medical School, London, UK (Prof K Dheda)

Correspondence to:

Prof Keertan Dheda, H46.41 Old Main Building, Groote Schuur Hospital Observatory, 7945 Cape Town, South Africa  
[keertan.dheda@uct.ac.za](mailto:keertan.dheda@uct.ac.za)

acquisition are urgently needed in primary care, where early diagnosis will have the greatest effect.<sup>8</sup>

Sputum can be safely acquired from sputum-scarce or smear-negative patients through induction using ultrasonic nebulisation with hypertonic saline.<sup>9–11</sup> Low cost, outdoor sputum induction booths with adequate infection control could help to make induction more feasible in resource-limited, HIV-prevalent primary care settings, and several already operate in South African primary care clinics. An alternative effective sputum-sampling method is instruction by a health-care worker,<sup>12,13</sup> in which a health-care worker provides simple training to the patient in sputum expectoration. However, which of these strategies is best for people with sputum-scarce or smear-negative tuberculosis in primary care is unclear. Furthermore, the effects of either technique on patient-oriented outcomes—eg, treatment initiation and time to treatment—in settings where empiric treatment is common, are unknown. Assessment of diagnostic strategies using patient-centred outcomes as primary endpoints is recognised by WHO advisory groups as essential for endorsement and scale-up.<sup>14</sup> Therefore, we did a randomised controlled trial to compare these two sampling strategies for adults with smear-negative or sputum-scarce tuberculosis in primary care.

See Online for appendix

## Methods

### Study design and participants

We did this open-label pragmatic randomised controlled trial in three primary care clinics in Cape Town, South Africa. The first patient was enrolled on Aug 7, 2009, and follow-up was completed on May 5, 2012. The study was approved by the University of Cape Town Human Research Ethics Committee.

Eligibility criteria were: age at least 18 years, ongoing symptoms suggestive of tuberculosis, and either an inability to self-expectorate a sputum sample or two negative sputum smear-microscopy samples (self-expectorated within the preceeding 4 weeks). We included both HIV-positive and HIV-negative patients. Patients were excluded if their initial spontaneous sputum samples were assessed with MTB/RIF assay rather than smear microscopy. Patients were compensated 50 rand for transport and absence from work when attending follow-up, non-routine, study clinic visits. Written informed consent was obtained from all patients and the standard of care was not altered by study participation.

### Randomisation and masking

Patients were referred for study screening by the designated nursing staff at each primary care clinic, after which they were assessed by a study research nurse. We used a simple randomisation strategy without stratification or masking. 600 unmarked, opaque envelopes each containing an intervention group assignment (in a 1:1 ratio) were made by personnel not involved in patient enrolment. The unmarked envelopes

were shuffled by hand and distributed to each clinic in batches of 100. At the clinic, enrolment was done by the study nurse before either a doctor's assessment or a chest radiograph. After providing informed consent and completing a detailed clinical record form, patients selected an envelope to determine their allocation. Intervention group cards were then stored with patient clinical record forms and frequent unannounced checks were made by the researcher to confirm adherence to the randomisation protocol.

### Procedures

Patients allocated to receive health-care worker instruction were individually instructed in their native language by the study nurse. Sputum induction was done by a trained study nurse using ultrasonically nebulised 5% hypertonic saline in an outdoor, open-air ventilated booth. Both instruction and induction occurred only once after enrolment. The appendix shows full step-by-step details of both procedures, as well as sample processing and laboratory methods.

All patients were asked to provide two spot sputum samples. Samples of at least 1 mL, irrespective of visual quality, were sent for processing at the National Health Laboratory (Cape Town, South Africa). Results for smear microscopy were available within 24 h. If a patient provided two samples they were randomly labelled sputum 1 and sputum 2. Sputum 1 was processed with N-acetyl-L-cysteine and sodium hydroxide, centrifuged, and resuspended in 1.5 mL phosphate buffer. The sample was subjected to auramine O staining and fluorescence microscopy; 0.5 mL of the sediment was inoculated into a Mycobacterial Growth Indicator Tube (Becton Dickinson Diagnostics; Franklin Lakes, NJ, USA) and incubated for no more than 8 weeks. Sputum 2 was unprocessed and frozen at  $-20^{\circ}\text{C}$  within 6 h of acquisition. Xpert MTB/RIF testing was unavailable at enrolment. After the study was completed, available sputum 2 specimens were thawed and tested with the Xpert MTB/RIF assay.<sup>15</sup> If patients provided only a single sputum sample, this was processed as sputum 1.

As per standard clinic guidelines, patients had chest radiography after enrolment and sputum sampling, and were scheduled to return to the clinic for a doctor's assessment. All patients—except those who were lost to follow-up or who had a positive sputum smear—were assessed by a doctor at least once, but usually twice; first, as soon as possible after enrolment (usually within 7 days), and second, at the 8 week follow-up visit (unless otherwise specified). Chest radiographs, treatment regimens, and culture results of all sputum smear-positive patients referred directly for treatment were reviewed by the study doctor. If extrapulmonary tuberculosis was suspected, additional non-sputum samples were taken at the doctor's discretion. The timing and initiation of treatment was decided by the attending doctor, and the basis for starting treatment (smear

microscopy, clinical or chest radiography [empiric], or culture) was recorded. Throughout the study, a specialist physician or pulmonologist reviewed the medical files of all patients who were treated, any patient of concern to the attending doctor, and a random selection of remaining, untreated patients. We calculated the cost of each sputum sampling strategy by an ingredients approach (total expenditure is presented as a sum of the components).<sup>16</sup> Costs are expressed in 2012 \$US at an exchange rate of \$1=ZAR8·20 based on the UN rate of exchange in September, 2012 (appendix).

The primary outcome was the proportion of patients who started treatment for tuberculosis during the 8 week study period. Secondary outcomes were: time-specific proportions of patients starting treatment within 3, 5, 7, 10, 14, 21, and 56 days from enrolment, the proportion of patients producing sputum for diagnostic testing, adverse effects related to sampling procedures, the quality of sputum samples as measured by the Bartlett score,<sup>17</sup> and tuberculosis case detection by diagnostic method (smear microscopy, MTB/RIF assay, or culture).

### Statistical analysis

Published<sup>9</sup> and unpublished data suggested that 15–20% of the study population would have a positive culture. Thus, we chose a target sample size of 500 patients, which would provide at least 80% power to

detect a 10% difference in the proportion of patients starting treatment (overall and at prespecified points), assuming roughly 15% treatment initiation in the instructed group and 25% in the induction group, with 5% type 1 error. We used STATA IC (version 10) for all statistical analyses. We did a modified intention-to-treat analysis with the  $\chi^2$  and Wilcoxon rank-sum tests to compare groups, with no corrections for multiple testing made for secondary outcomes. We calculated point estimates and odds ratios (ORs) with 95% CIs together with p values, all of which were two-sided.

The study is registered with ClinicalTrials.gov, number NCT01545661.

### Role of the funding source

The sponsor of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

### Results

Figure 1 shows the trial profile. We screened 517 patients and included 481 in the analysis (213 assigned to health-care worker instruction, 268 assigned to sputum induction). Table 1 shows baseline characteristics.

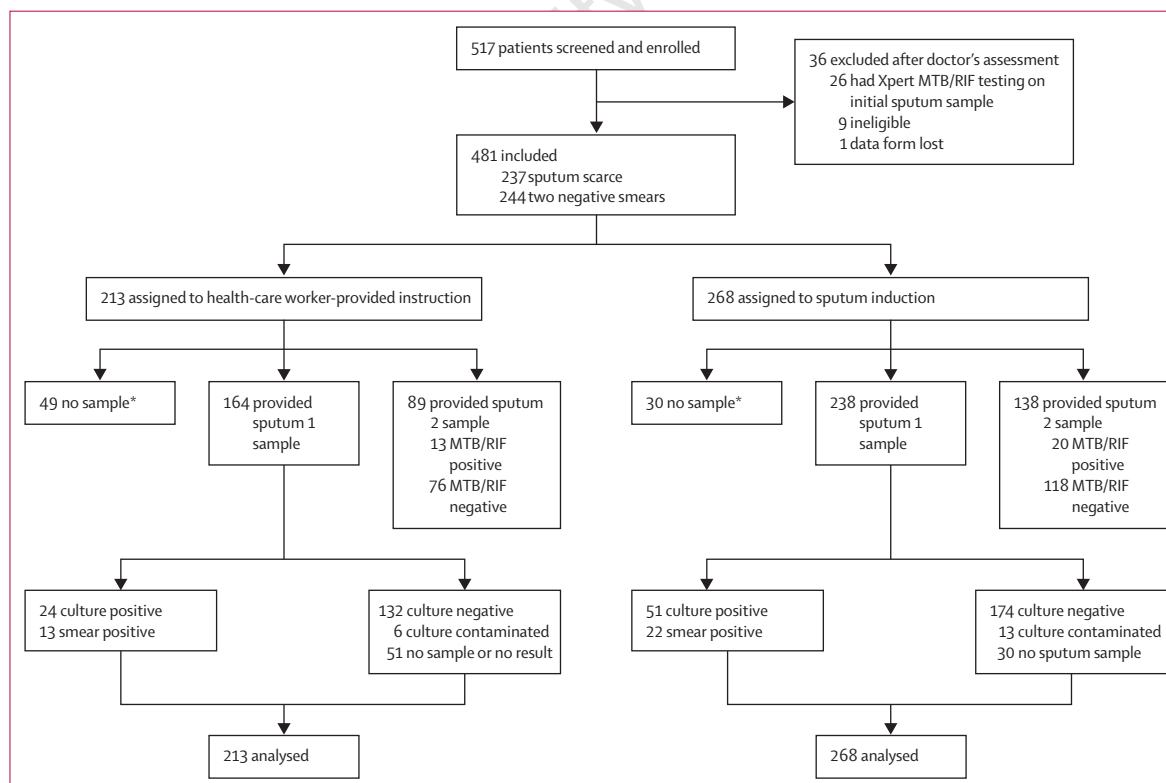


Figure 1: Trial profile

\*Unable to provide a sputum sample for diagnostic testing.

262 patients were male and 171 were HIV positive. In patients who were HIV positive at enrolment, median CD4 cell count was 242 cells per mL (IQR 146–358) and 37 of 171 (22%) were receiving antiretroviral therapy. Baseline characteristics did not differ substantially between groups. Cough duration and phlegm

	Health-care worker instruction (n=213)	Induction (n=268)
Age (years)	40 (31–49)	38 (29–49)
Men	122 (57%)	140 (52%)
HIV-positive	75 (35%)	96 (36%)
CD4 cell count (cells per mL)	239 (136–345)	247 (149–379)
Taking antiretrovirals	15 (20%)	22 (23%)
History of tuberculosis	82 (38%)	98 (37%)
Diagnostic categorisation		
Two negative sputum smears	117 (55%)	127 (47%)
Unable to produce sputum	96 (45%)	141 (53%)
Cough >2 weeks	189 (89%)	241 (90%)
Productive cough	141 (66%)	170 (63%)
Night sweats	152 (71%)	192 (72%)
Weight loss	145 (68%)	190 (71%)
Appetite loss	114 (54%)	139 (52%)
Bodyweight (kg)	62 (54–72)	63 (55–72)
Chest radiography compatible with tuberculosis	85 (40%)	94 (35%)

Data are median (IQR) or n (%).

**Table 1: Baseline characteristics**

production were the only differences between patients producing one and two sputum samples (appendix).

53 of 213 (25%) patients who had health-care worker instruction versus 73 of 268 (27%) who had induction started treatment by week 8 (OR 0·88, 95% CI 0·57–1·36;  $p=0\cdot56$ ; table 2, figure 2A). At 3, 5, 7, 14, 21, and 56 days after enrolment the proportion of patients in each group who had started treatment did not differ significantly (figure 2B). At 10 days—the median time of the doctor’s first visit—40 (75%) of 53 instructed patients compared with 43 (59%) of 73 induced patients had started treatment ( $p=0\cdot053$ ). The median time to start of treatment was shorter for instructed versus induced patients (4 days, IQR 2–9 vs 7 days, IQR 2–27;  $p=0\cdot029$ ). However, neither the proportion of patients who started treatment by a specific time nor median times to treatment differed significantly between groups if the analysis included only patients in whom a sputum sample was acquired for diagnostic testing (appendix).

Irrespective of intervention group, patients unable to produce a sputum sample, and thus not waiting for a diagnostic test result, received a doctor’s assessment and empiric treatment quicker than did those who could produce a sample (median 3 days, IQR 1–7 vs 6 days, 2–8 days;  $p=0\cdot004$ ). Furthermore, if the analysis was restricted to sputum-scarce or HIV-positive patients, or repeated with a random sample that had balanced patient numbers from each study group ( $n=200$ ; appendix) time-specific proportions of patients starting treatment did not differ between groups. We also did a secondary time-to-event analysis, comparing the time to start of treatment between groups (appendix). Overall ( $p=0\cdot4$ ), and when the analysis was restricted to patients who definitely had tuberculosis ( $p=0\cdot7$ ), the groups did not differ significantly.

In our analysis by reason for starting treatment, a similar proportion of patients started treatment in the instructed group versus the induced group, whether treatment initiation was based on positive sputum smear microscopy or positive culture (table 2). By contrast, more instructed patients compared with induced patients received treatment empirically based on clinical and radiological findings (table 2). 60% of empirically treated patients were HIV positive and treatment was started in accordance with the 2007 WHO guidelines.<sup>18</sup>

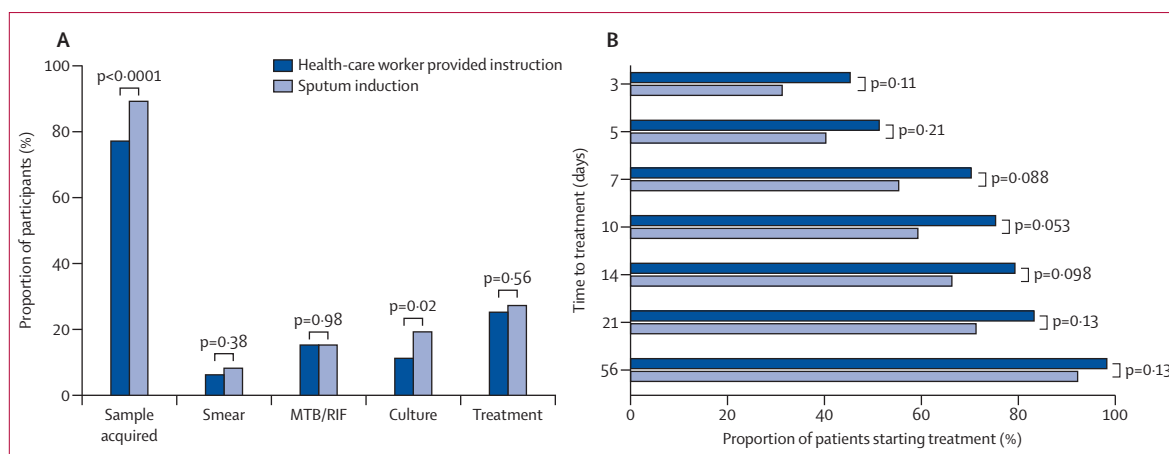
A smaller proportion of instructed patients compared with induced patients successfully produced a sputum sample of at least 1 mL for diagnostic testing (164/213 [77%] vs 238/268 [89%];  $p<0\cdot0001$ ; figure 2). However, the proportion of samples that were of good quality—as assessed by the Bartlett score—was much the same between groups (table 3).

27 induced patients (10%) and 31 (15%) instructed patients who provided a sputum specimen did not return for an initial doctor’s assessment or to collect their diagnostic test results ( $p=0\cdot1$ ). After 2 months, 12 of the initial 27 induced patients (4/12 were culture positive) and 13 of the 31 instructed patients (3/13 were culture

	Health-care worker instruction	Induction	OR (95% CI)	p value
Total patients starting treatment	53/213 (25%)	73/268 (27%)	0·88 (0·57–1·36)	0·56
Diagnosis and treatment initiation based on smear microscopy				
Patients treated	13/53 (25%)	22/73 (30%)	0·75 (0·31–1·80)	0·49
Median time to treatment (days)	2 (2–6)	3 (2–6)	0·03 (0·00–27·15)	0·68
Diagnosis based on clinical and radiological presentation with empiric treatment initiation				
Patients treated	32/53 (60%)	28/73 (38%)	2·45 (1·12–5·40)	0·015
Patients with HIV*	18/32 (56%)	17/28 (61%)	0·83 (0·26–2·57)	0·73
Patients without HIV	14/32 (44%)	11/28 (39%)	1·20 (0·38–3·83)	0·73
Median time to treatment (days)	4 (1–9)	7 (3–10)	0·01 (0·00–4·01)	0·15
Diagnosis and treatment initiation based on culture				
Patients treated based on sputum 1 culture result	6/53 (11%)	18/73 (25%)	0·39 (0·12–1·14)	0·060
Patients treated based on other (sputum 2 or non-sputum) culture results	1/53 (2%)	2/73 (3%)	0·68 (0·01–13·47)	0·76
Median time to treatment (days)	34 (29–48)	42 (20–56)	0·26 (0·00–5·48)	0·90
Culture-positive patients not given any tuberculosis treatment during study	3/213 (1%)	4/268 (2%)	0·94 (0·14–5·61)	0·93

One patient in the health-care worker instruction group and three in the sputum induction group were missing data for reason for treatment initiation. \*Initiation of tuberculosis treatment was based on the 2007 WHO smear-negative diagnostic and treatment algorithm for HIV-positive ambulatory patients.

**Table 2: Outcomes stratified by method of diagnosis**



**Figure 2: Proportions of patients receiving tuberculosis diagnosis and treatment and time to treatment initiation** (A) shows proportions of samples acquired, diagnosis by smear microscopy, Xpert MTB/RIF assay, and culture, and proportions of patients given treatment. (B) shows time to start of treatment. Groups were compared with  $\chi^2$ .

	Instruction by health-care worker (n=213)	Sputum induction (n=268)	p value
<b>Sample volume and quality</b>			
Patients providing at least one sputum sample of $\geq 1$ mL for laboratory testing*	164 (77%)	238 (89%)	<0.0001
Sputum samples considered to be adequate quality	117/136 (86%)	181/216 (84%)	0.57
Side-effects from sampling procedure	0 (0%)	32 (12%)†	<0.0001
<b>Diagnostic yield and accuracy</b>			
Smear microscopy yield	13 (6%)	22 (8%)	0.38
Tuberculosis culture yield	24 (11%)	51 (19%)	0.020
Median time to positivity for tuberculosis culture (IQR; days)	14 (11–18)	13 (9–18)	0.54
MTB/RIF assay diagnostic yield (sputum sample 2)	13/89 (15%)	20/138 (14%)	0.98
MTB/RIF assay sensitivity‡			
All culture positive sensitivity (n/N; %; 95% CI)	9/12 (75%; 51–99)	16/25 (64%; 45–83)	0.50
Sputum culture positive sensitivity (n/N; %; 95% CI)	4/7 (57%; 20–94)	8/16 (50%; 26–75)	0.75

\*19 sputum specimens undergoing liquid culture were contaminated: six in the health-care worker-provided instruction group and 13 in the sputum induction group.  
†Includes nausea or vomiting, headache, dizziness, and shortness of breath. ‡Calculated using liquid tuberculosis culture from a paired sample as the reference standard.

**Table 3: Diagnostic outcomes**

positive) were still lost to follow-up ( $p=0.8$ ). Diagnostic yield from smear microscopy was similar in instructed and induced patients (table 3), as was diagnostic yield from MTB/RIF assay (table 3). In view of the overall similarities between patients producing one and two sputum samples (appendix), we calculated an estimated MTB/RIF diagnostic yield for sputum 1 (ie, adjusting for the success of sample acquisition), which provided much the same diagnostic yield in instructed patients (24/213 [11%]) versus induced patients (43/268 [16%];  $p=0.1$ ). By contrast, culture-based diagnostic yield was lower in instructed patients compared with induced patients (table 3), although culture-based diagnostic yield did not differ significantly if analysis was restricted to patients providing a sputum sample for diagnostic testing (23/164 [14%] vs 51/238 [21%];  $p=0.060$ ). Among

culture-positive patients, median time to culture positivity was similar in instructed and induced patients (table 3).

Side-effects related to sample acquisition were reported in 32 of 268 (12%) patients who had sputum induction and none who had health-care worker instruction (table 3). The most common side-effects were: shortness of breath ( $n=11$ ), dizziness ( $n=9$ ), headache ( $n=8$ ), and nausea or vomiting ( $n=7$ ). Sputum induction was stopped if patients had side-effects and all side-effects resolved without the need for review by a doctor.

Health-care worker-provided instruction cost \$2.14 per sampling procedure versus \$7.88 for sputum induction. The higher cost of sputum induction is a result of the additional consumables used and staff time needed for nebulisation (appendix).

## Discussion

To our knowledge, this study is the first pragmatic randomised controlled trial to compare health-care worker instruction with induction for sputum sampling in adults with suspected tuberculosis who are smear-negative or sputum-scarce in a primary care practice (panel). Our findings have important clinical and public health policy implications. In regions where HIV is common, smear-negative and sputum-scarce tuberculosis presents a diagnostic challenge. Our study supports the use of health-care worker instruction as the initial sputum sampling strategy. Nurses should instruct patients how to take a sample before patients are empirically treated or given sputum induction. Although sputum sampling by induction provided an adequate specimen volume and microbiological diagnosis in a higher proportion of patients than did health-care worker instruction, it was more costly and did not result in more patients starting treatment. Notably, sputum induction did not result in a higher proportion of case detection using same-day diagnostic methods (smear microscopy and Xpert MTB/RIF), probably because of the paucibacillary nature of induced sputum. This result, combined with the high rates of empiric treatment initiation based on clinical and radiological findings,

meant that the benefits of sputum induction failed to affect either the proportion of patients starting treatment or the time to start of treatment. Thus, health-care worker instruction had an equivalent effect on treatment initiation compared with induction, at a substantially lower cost and with fewer adverse events.

In previous studies, health-care worker instruction<sup>12,13</sup> and nurse-specific educational outreach improved rates of tuberculosis case detection with smear microscopy in people with persistent cough in primary care.<sup>26,27</sup> In HIV-positive Malawian patients thought to have tuberculosis, health-care worker instruction offered better diagnostic yield for smear microscopy and culture than did alternative acquisition methods.<sup>22</sup> However, before this study, no comparative randomised controlled studies of sampling strategies were available and despite its simplicity, health-care worker instruction is neither routinely used nor is it a formalised step in smear-negative tuberculosis diagnostic procedures. Our study findings suggest that national tuberculosis programmes should include health-care worker instruction as the first strategy for smear-negative or sputum-scarce patients and thus, they should urgently provide widespread training to health-care workers and nurses about this sputum sampling strategy.

Other studies<sup>9,11,20</sup> have shown sputum induction to be an excellent and safe sampling method for culture-based diagnosis. Although our study does not change this conclusion, we have found that sputum induction does not necessarily affect treatment initiation because of the long delays associated with culture-based diagnosis. Furthermore, because Xpert MTB/RIF assay performs suboptimally when using induced sputum specimens, the use of Xpert MTB/RIF assay as a replacement for smear microscopy would probably not have affected the primary outcome. Thus, sputum induction has limitations for adults with smear-negative and sputum-scarce tuberculosis, particularly in settings where other investigations—eg, chest radiography and high empiric treatment use—are routinely done.

Although not assessed in this study, a step-wise approach might be best for diagnosis of smear-negative or sputum-scarce patients, with routine use of sputum induction reserved for when instruction has been unsuccessful or when a culture-based diagnosis is essential—eg, in a suspected case of multidrug-resistant tuberculosis.<sup>28,29</sup> More studies are needed to assess such an approach. In addition, sputum induction is still an important sampling strategy for children and asymptomatic HIV-positive patients who are being screened for tuberculosis before starting antiretroviral therapy.<sup>21,30</sup>

Our study has some limitations. An open-label design can be prone to bias but this was chosen for its simplicity because of the location and infrastructure of the clinics and the nature of the intervention. However, we did regular unannounced checks—to ensure that the

### Panel: Research in context

#### Systematic review

We searched PubMed for studies about either sputum induction or health-care worker instruction published in English up to March 13, 2013. We combined search terms that could indicate sputum induction or health-care worker instruction (“sputum induction, induced sputum, sputum expect\*, sputum sampl\*, sputum/\*microbiology”) with “TB”. We identified two systematic reviews of sputum induction<sup>9,11</sup> and a large study<sup>19</sup> of adult patients with suspected smear-negative and sputum-scarce tuberculosis. We identified four studies<sup>12,13,20,21</sup> involving health-care worker supervision or instruction during sputum sampling.

#### Interpretation

To our knowledge, our study is the only randomised controlled trial to assess the role of sputum induction in the diagnosis of tuberculosis, and is the first study to directly compare two sputum sampling strategies with treatment uptake as the primary outcome. Previous studies of sputum induction in adults with suspected smear-negative and sputum-scarce tuberculosis from settings with high HIV and tuberculosis prevalence are heterogeneous, with varying estimates of culture-based diagnostic yield (8–66%) and smear microscopy sensitivity (32–60%).<sup>19,22–25</sup> Two studies of health-care-provided instruction show increased diagnosis of tuberculosis by smear for adults with suspected pulmonary tuberculosis,<sup>12,13</sup> but two studies using sequential combinations of instruction and induction provided conflicting results.<sup>20,22</sup> Our study confirms that sputum induction—although more costly—offers better sputum sampling and increased culture-based diagnosis compared with simple instruction. However, use of sputum induction did not result in more patients being treated. Health-care worker-provided instruction should be the preferred initial sputum sampling strategy for adults with suspected smear-negative and sputum-scarce tuberculosis, especially where sputum induction facilities are unavailable or where culture-based diagnosis is unlikely to alter treatment decision making. Advocacy to improve the training of health-care workers in sputum sampling instruction and the incorporation of simple instruction into diagnostic algorithms of primary care clinics is warranted.

protocol was adhered to—and patient characteristics did not differ between groups, suggesting no bias. The different number of patients randomly assigned to each group—although statistically plausible given the simple randomisation method—might have introduced selection bias. However, baseline characteristics were similar and the main conclusions were the same when we analysed a random sample of 200 patients with balanced groups. The exclusion of 36 patients from the primary analysis is another important limitation, with most excluded because of programmatic implementation of Xpert MTB/RIF instead of smear microscopy for testing pre-enrolment sputum specimens at some study sites in the final few months of enrolment. Sensitivity analyses showed no differences between excluded and included patients, and power calculations suggest that the analysis had a greater than 80% power to detect a 12% difference between study groups for the primary outcome taking exclusions into account.

No validated sputum quality scoring system exists for induced sputum samples. Thus, Bartlett scoring is not ideal and conclusions about differences in sample quality between groups should be interpreted with caution. Empiric treatment was more common among instructed patients and whether this constituted appropriate treatment or over-treatment is difficult to ascertain. The exact specificity of empiric treatment is unknown, and estimates from studies of WHO algorithms for smear-negative tuberculosis in high tuberculosis and HIV settings range from 44% to 95%.<sup>31–34</sup> In our study, 60% of empirically treated patients were HIV-positive and qualified for treatment in accordance with the WHO smear-negative tuberculosis algorithm.<sup>22</sup> Xpert MTB/RIF assay was used on stored sputum samples when available and not for treatment decisions. Because many patients did not have a second sputum specimen these findings should be interpreted cautiously. Our findings are applicable to settings in which HIV is common and further studies are needed to assess their usefulness elsewhere.

Our data support the use of nurse-driven health-care worker instruction as the initial sputum sampling method for adults with suspected smear-negative or sputum-scarce tuberculosis in a high-burden primary care setting. Sputum induction is an important sampling strategy when the need for a microbiologically confirmed diagnosis of tuberculosis is essential. More effort should be made to formalise and incorporate sputum instruction and supervision in the education of primary clinic health-care workers in regions where HIV and tuberculosis are common.

#### Contributors

JGP and KD designed the study. JGP, JT, and MP collected data. JP, GT, and AP analysed the data. JGP wrote the first draft of the report and all authors gave input to the final version.

#### Conflicts of interest

We declare that we have no conflicts of interest.

#### Acknowledgments

This work was supported by the South African National Research Foundation, a TBSugent grant from the European Commission, the National Institutes of Health, the European and Developing Countries Clinical Trials Partnership, and the Discovery Foundation. We thank the research nursing staff who were involved in recruitment of patients and collection of sputum samples. In addition, the assistance and support of the health-care workers at the primary care clinics of Langa, Gugulethu, and Chapel Street was greatly appreciated.

#### References

- 1 WHO. WHO Report 2011: global tuberculosis control. Geneva: World Health Organization, 2011.
- 2 Lonnroth K, Castro KG, Chakaya JM, et al. Tuberculosis control and elimination 2010–50: cure, care, and social development. *Lancet* 2010; **375**: 1814–29.
- 3 Getahun H, Harrington M, O'Brien R, Nunn P. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: informing urgent policy changes. *Lancet* 2007; **369**: 2042–49.
- 4 WHO. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. Geneva: World Health Organization, 2011.
- 5 Boehme CC, Nicol MP, Nabeta P, et al. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 2011; **377**: 1495–505.
- 6 Theron G, Peter J, Dheda K. Xpert MTB/RIF test for tuberculosis. *Lancet* 2011; **378**: 481.
- 7 Theron G, Pooran A, Peter J, et al. Do adjunct TB tests, when combined with Xpert MTB/RIF, improve accuracy and the cost of diagnosis in a resource-poor setting? *Eur Respir J* 2012; **40**: 161–68.
- 8 Dye C, Williams BG. The population dynamics and control of tuberculosis. *Science* 2010; **328**: 856–61.
- 9 Hepple P, Ford N, McEnerney R. Microscopy compared to culture for the diagnosis of tuberculosis in induced sputum samples: a systematic review. *Int J Tuberc Lung Dis* 2012; **16**: 579–88.
- 10 Schoch OD, Rieder P, Tueller C, et al. Diagnostic yield of sputum, induced sputum, and bronchoscopy after radiologic tuberculosis screening. *Am J Respir Crit Care Med* 2007; **175**: 80–86.
- 11 Gonzalez-Angulo Y, Wiysonge CS, Geldenhuys H, et al. Sputum induction for the diagnosis of pulmonary tuberculosis: a systematic review and meta-analysis. *Eur J Clin Microbiol Infect Dis* 2012; **31**: 1619–30.
- 12 Alisjahbana B, van Crevel R, Danusantoso H, et al. Better patient instruction for sputum sampling can improve microscopic tuberculosis diagnosis. *Int J Tuberc Lung Dis* 2005; **9**: 814–17.
- 13 Khan MS, Dar O, Sismanidis C, Shah K, Godfrey-Faussett P. Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial. *Lancet* 2007; **369**: 1955–60.
- 14 Cobelens F, van den Hof S, Pai M, Squire BS, Ramsay A, Kimmmerling ME. Which new diagnostics for tuberculosis, and when? *J Infect Dis* 2012; **205** (suppl 2): S191–98.
- 15 Boehme CC, Nabeta P, Hillemann D, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med* 2010; **363**: 1005–15.
- 16 WHO. Making choices in health: WHO guide to cost-effectiveness analysis. Geneva: World Health Organization, 2003.
- 17 Bartlett RC. Medical microbiology: quality, cost and clinical relevance. New York: John Wiley and Sons, 1974.
- 18 WHO. Improving the diagnosis and treatment of smear-negative pulmonary and extrapulmonary tuberculosis among adults and adolescents. Geneva: World Health Organization, 2006.
- 19 Peter JG, Theron G, Singh N, Singh A, Dheda K. Sputum induction to aid the diagnosis of smear-negative or sputum-scarce TB in adults from a HIV-endemic setting. *Eur Respir J* 2013; published online March 21. DOI:10.1183/09031936.00198012.
- 20 Chang KC, Leung CC, Yew WW, Tam CM. Supervised and induced sputum among patients with smear-negative pulmonary tuberculosis. *Eur Respir J* 2008; **31**: 1085–90.

- 21 Lawn SD, Brooks SV, Kranzer K, et al. Screening for HIV-associated tuberculosis and rifampicin resistance before antiretroviral therapy using the Xpert MTB/RIF assay: a prospective study. *PLoS Med* 2011; **8**: e1001067.
- 22 Bell DJ, Dacombe R, Graham SM, et al. Simple measures are as effective as invasive techniques in the diagnosis of pulmonary tuberculosis in Malawi. *Int J Tuberc Lung Dis* 2009; **13**: 99–104.
- 23 Wilson D, Nachega J, Morroni C, Chaisson R, Maartens G. Diagnosing smear-negative tuberculosis using case definitions and treatment response in HIV-infected adults. *Int J Tuberc Lung Dis* 2006; **10**: 31–38.
- 24 Morse M, Kessler J, Albrecht S, et al. Induced sputum improves the diagnosis of pulmonary tuberculosis in hospitalized patients in Gaborone, Botswana. *Int J Tuberc Lung Dis* 2008; **12**: 1279–85.
- 25 Parry CM, Kamoto O, Harries AD, et al. The use of sputum induction for establishing a diagnosis in patients with suspected pulmonary tuberculosis in Malawi. *Tuberc Lung Dis* 1995; **76**: 72–76.
- 26 Fairall L, Bachmann MO, Zwarenstein M, et al. Cost-effectiveness of educational outreach to primary care nurses to increase tuberculosis case detection and improve respiratory care: economic evaluation alongside a randomised trial. *Trop Med Int Health* 2010; **15**: 277–86.
- 27 Fairall LR, Zwarenstein M, Bateman ED, et al. Effect of educational outreach to nurses on tuberculosis case detection and primary care of respiratory illness: pragmatic cluster randomised controlled trial. *BMJ* 2005; **331**: 750–54.
- 28 Dheda K, Warren RM, Zumla A, Grobusch MP. Extensively drug-resistant tuberculosis: epidemiology and management challenges. *Infect Dis Clin North Am* 2010; **24**: 705–25.
- 29 Dheda K, Shean K, Zumla A, et al. Early treatment outcomes and HIV status of patients with extensively drug-resistant tuberculosis in South Africa: a retrospective cohort study. *Lancet* 2010; **375**: 1798–807.
- 30 Nicol MP, Zar HJ. New specimens and laboratory diagnostics for childhood pulmonary TB: progress and prospects. *Paediatr Respir Rev* 2011; **12**: 16–21.
- 31 Alamo ST, Kunutsor S, Walley J, et al. Performance of the new WHO diagnostic algorithm for smear-negative pulmonary tuberculosis in HIV prevalent settings: a multisite study in Uganda. *Trop Med Int Health* 2012; **17**: 884–95.
- 32 Huerga H, Varaine F, Okwaro E, et al. Performance of the 2007 WHO algorithm to diagnose smear-negative pulmonary tuberculosis in a HIV prevalent setting. *PLoS One* 2012; **7**: e51336.
- 33 Soto A, Solari L, Gotuzzo E, Acinelli R, Vargas D, Van der Stuyf P. Performance of an algorithm based on WHO recommendations for the diagnosis of smear-negative pulmonary tuberculosis in patients without HIV infection. *Trop Med Int Health* 2011; **16**: 424–30.
- 34 Wilson D, Mbhele L, Badri M, et al. Evaluation of the World Health Organization algorithm for the diagnosis of HIV-associated sputum smear-negative tuberculosis. *Int J Tuberc Lung Dis* 2011; **15**: 919–24.

## Chapter 5.

### **Evaluation of the Xpert® MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting**

Theron, G.\*, Peter, J.G.\*, van Zyl-Smit, R. N., Mishra, H., Streicher, E., Murray, S, Dawson, R., Whitelaw, A., Hoelscher, A., Sharma, S., Pai, M., Warren, R., and Dheda, K. <sup>1</sup>

\*The first two authors contributed equally.

Published: *Am J of Resp Crit Care Med* 2011 Jul 1; 184(1):132-40

#### **PhD Context**

In December 2010, the World Health Organisation endorsed the use of the MTB/RIF assay, making a strong recommendation for its use as the frontline TB diagnostic in HIV-infected and previous TB patients at risk for MDR-TB (1). Shortly thereafter, in March 2011, the SA Minister of Health Dr Aaron Motsoaledi, announced that the Department of Health would replace smear-microscopy with MTB/RIF for all frontline TB diagnosis with a phased implementation commencing immediately. This WHO endorsement and recommendation was based almost entirely on the Foundation of Innovative and Novel Diagnostics (FIND)'s demonstration study findings and evaluation study preliminary findings (2, 3). However, relatively few smear-negative cases from a high HIV prevalence setting were included in these large studies, and in addition many unanswered research questions about MTB/RIF performance remained. This study was conducted and is presented as part of the thesis because i) the MTB/RIF assay, being used for frontline TB diagnosis, may significantly improve

the diagnosis of smear-negative culture-positive TB, and ii) many novel research questions pertaining to the use of MTB/RIF had not been previously investigated.

University of Cape Town

# Evaluation of the Xpert MTB/RIF Assay for the Diagnosis of Pulmonary Tuberculosis in a High HIV Prevalence Setting

Grant Theron<sup>1\*</sup>, Jonny Peter<sup>1,\*</sup>, Richard van Zyl-Smit<sup>1</sup>, Hridesh Mishra<sup>2</sup>, Elizabeth Streicher<sup>3</sup>, Samuel Murray<sup>1</sup>, Rodney Dawson<sup>1</sup>, Andrew Whitelaw<sup>4</sup>, Michael Hoelscher<sup>5</sup>, Surendra Sharma<sup>2</sup>, Madhukar Pai<sup>6</sup>, Robin Warren<sup>3</sup>, and Keertan Dheda<sup>1,7,8</sup>

<sup>1</sup>Lung Infection and Immunity Unit, Division of Pulmonology and UCT Lung Institute, Department of Medicine, <sup>4</sup>Division of Medical Microbiology, and <sup>7</sup>Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa; <sup>2</sup>Department of Medicine, All India Institute of Medical Sciences, New Delhi, India; <sup>3</sup>DST/NRF Centre of Excellence for Biomedical TB Research/MRC Centre for Molecular and Cellular Biology, Stellenbosch University, Stellenbosch, South Africa; <sup>5</sup>Department for Infectious Diseases and Tropical Medicine, Klinikum of the University of Munich, Munich, Germany; <sup>6</sup>Department of Epidemiology and Biostatistics, McGill University, Montreal, Canada; and <sup>8</sup>Department of Infection, University College London Medical School, London, United Kingdom

**Rationale:** Xpert MTB/RIF is a novel automated molecular diagnostic recently endorsed by the World Health Organization. However, performance-related data from high HIV prevalence settings are limited.

**Objectives:** The impact of sample-related factors on performance and the significance of Xpert MTB/RIF-positive culture-negative discordance remain unclear.

**Methods:** Xpert MTB/RIF was evaluated using single archived spot-sputum samples from 496 South African patients with suspected TB. *Mycobacterium tuberculosis* culture positivity and phenotypic resistance to rifampicin served as reference standards.

**Measurements and Main Results:** Overall, Xpert MTB/RIF detected 95% (95% confidence interval [CI], 88–98%; 89 of 94) of smear-positive culture-positive cases and the specificity was 94% (91–96%; 320 of 339). The sensitivity in smear-negative cases was 55% (35–73%; 12 of 22) when the analysis was restricted to 1 ml of unprocessed sputum and culture time-to-positivity of less than or equal to 28 days. Compared with smear microscopy (n = 94), Xpert MTB/RIF detected an additional 17 cases (n = 111) representing an 18% (11–27%; 111 vs. 94) relative increase in the rapid TB case detection rate. Moreover, compared with smear microscopy, the inclusion of Xpert MTB/RIF-positive culture-negative TB cases (ruled-in by an alternative diagnostic method) resulted in the detection of a further 16 cases (n = 127), thus significantly increasing the rapid TB case detection rate to 35% (95% CI, 26–45%; 94 to 111 vs. 94 to 127;  $P < 0.01$ ), the overall specificity to 99.1% (97–100%; 320 of 323;  $P < 0.001$ ), and sensitivity in smear-negative TB to 60% ( $P = 0.12$ ). Performance strongly correlated with smear status and culture time-to-positivity. In patients infected with HIV compared with patients uninfected with HIV Xpert MTB/RIF showed a trend to reduced sensitivity ( $P = 0.09$ ) and significantly reduced negative predictive value ( $P = 0.01$ ). The negative predictive value for rifampicin resistance was 99.4%.

(Received in original form January 12, 2011; accepted in final form April 14, 2011)

Supported by the Foundation for Innovative New Diagnostics, SA MRC (K.D.), SA DST SARChI (K.D.), EU-FP7 and EDCTP (K.D., G.T., J.P., R.V.Z.S., and M.P.), and the Canadian Institute for Health Research (M.P. and K.D.).

**Author contributions:** G.T., J.P., R.W., and K.D. designed the study. G.T., E.S., H.M., and S.M. generated the data. G.T., J.P., and K.D. wrote the first draft. G.T., K.D., J.P., R.V.Z.S., R.D., A.W., M.H., S.S., M.P., and R.W. analyzed the data and revised the manuscript.

\*Equally contributing joint first authors.

Correspondence and requests for reprints should be addressed to Keertan Dheda, M.B.Ch.B., Ph.D., H47 Old Main Building, Groote Schuur Hospital, Observatory, 7925, Cape Town, South Africa. E-mail: keertan.dheda@uct.ac.za

This article has an online supplement, which is accessible from this issue's table of contents at [www.atsjournals.org](http://www.atsjournals.org)

Am J Respir Crit Care Med Vol 184, pp 132–140, 2011

Originally Published in Press as DOI: 10.1164/rccm.201101-0056OC on April 14, 2011  
Internet address: [www.atsjournals.org](http://www.atsjournals.org)

## AT A GLANCE COMMENTARY

### Scientific Knowledge on the Subject

Xpert MTB/RIF is an accurate rapid diagnostic tool for tuberculosis (TB) and rifampicin resistance. Although it has been recently endorsed by the World Health Organization, there are limited data about the impact of HIV coinfection and sample-related factors on test performance. The significance of Xpert MTB/RIF-positive culture-negative samples remains unclear.

### What This Study Adds to the Field

HIV coinfection, but likely not sputum volume and processing methods, may impact on assay performance. Almost all Xpert MTB/RIF-positive culture-negative samples are likely true positives and this approximately doubles the number of detected TB cases over and above that of smear microscopy.

**Conclusions:** Xpert MTB/RIF outperformed smear microscopy, established a diagnosis in a significant proportion of patients with smear-negative TB, detected many highly likely TB cases missed by culture, and accurately ruled out rifampicin-resistant TB. Sample-specific factors had limited impact on performance. Performance in patients infected with HIV, especially those with advanced immunosuppression, warrants further study.

**Keywords:** smear-negative tuberculosis; tuberculosis; diagnostics; HIV; PCR

Tuberculosis (TB) is a major global health priority and kills approximately 1.7 million people annually (1). The incidence of multidrug resistant (MDR) TB is increasing with almost 0.5 million estimated new cases in 2008 (2). Although smear microscopy is widely used for the rapid diagnosis of TB, it does not detect drug resistance and sensitivity in individuals coinfecting with HIV varies between 20% and 50% (3). Results of mycobacterial culture often only become available after 2–8 weeks (4). This creates a diagnostic delay that hampers disease control, enhances transmission, and increases healthcare costs (5).

Xpert MTB/RIF (Cepheid, Sunnyvale, CA) is an automated user-friendly real-time polymerase chain reaction (PCR) assay designed for the rapid and simultaneous detection of *Mycobacterium tuberculosis* and rifampicin resistance (6–8). The assay amplifies a *M. tuberculosis* complex-specific region of the *rpoB* gene, which is probed with molecular beacons to detect the

presence of rifampicin resistance-determining mutations (9). In December 2010, the World Health Organization (WHO) endorsed the scale-up of Xpert MTB/RIF and recommended its use as the initial test in patients coinfected with HIV and TB and patients with suspected MDR TB (10, 11). The performance of the test with the first-generation software using both NALC-NaOH decontaminated and unprocessed sputum from 1,730 patients with suspected TB was recently assessed as part of a large multicentre study (7). Using a single assay on a single unprocessed sputum sample *M. tuberculosis* complex-specific DNA was detected in 98% of smear-positive cases and in 72% of smear-negative cases using culture positivity as a reference standard. However, only 67 (39%) of all smear-negative cases were from a high HIV prevalence setting. Moreover, there is no information regarding Xpert MTB/RIF performance stratified by CD4 count. Thus, data about performance in patients infected with HIV, particularly those with smear-negative TB, are limited.

There are several other gaps in the knowledge. (1) There are limited data about the effects of an alteration in sample volume (using the recommended 1 ml vs. <1 ml) or processing methods (raw vs. liquefied sputum) on assay performance. Additionally, the relationship between bacterial load (measured using smear grade and culture time-to-positivity [TTP]) and assay performance is unclear. These factors have important implications for data interpretation in sputum-scarce patients, the integration of the assay into existing laboratory work flows, and the design of future clinical trials. (2) What additional yield Xpert MTB/RIF can offer, if any, over culture is unknown. Thus, the significance of Xpert MTB/RIF-positive, culture-negative samples remains unclear. (3) The impact, given resource constraints, of combining smear microscopy and Xpert MTB/RIF requires clarification. (4) Finally, we evaluated, hitherto untested, the specificity of a recently released second-generation software algorithm for the simultaneous detection of *M. tuberculosis* and rifampicin resistance.

The goal was to validate test performance further using a single cartridge in the context of high HIV prevalence and to assess the impact of the previously mentioned factors on assay performance. These issues gain importance as countries prepare to roll-out and scale-up the Xpert MTB/RIF assay (10).

## METHODS

### Study Sites and Population

Sputa were collected from 496 consecutively recruited ambulant patients with suspected TB ( $\geq 18$  yr of age) between February 2007 and April 2010 at two primary care clinics in Cape Town, South Africa. Informed consent was obtained from all participants and the study was approved by the University of Cape Town, Faculty of Health Sciences Research Ethics Committee. Detailed patient and laboratory-specific information were recorded on a standardized case record form and captured using double data entry. An HIV test was performed after appropriate counseling. All chest radiographs were independently scored by two trained readers using the Chest Radiographic Reading and Reporting System (12, 13). Chest radiographs were scored as compatible or unlikely to be compatible with active TB. Discrepant results were adjudicated by a third senior reader.

### TB Case Definitions

Each patient was allocated to one of three diagnostic categories: (1) definite TB: a clinical presentation compatible with TB with at least one spot sputum sample culture-positive for *M. tuberculosis*; (2) probable TB: a clinical–radiologic picture highly suggestive of TB or anti-TB treatment was initiated by an attending clinician based on clinical suspicion but the patient did not meet the criteria for definite TB (no culture-based evidence of *M. tuberculosis*); or (3) non-TB: no evidence

of TB based on smear microscopy and culture, no anti-TB treatment initiated with response to alternative treatment where appropriate, and when available no radiologic evidence to support the diagnosis of TB.

### Microbiology

At the first visit two paired spot sputa were concurrently collected from each patient. One arbitrarily selected sample was decontaminated in NALC-NaOH, submitted for routine concentrated fluorescence smear microscopy, and cultured for *M. tuberculosis* using the BACTEC MGIT 960 system (BD Diagnostics, Franklin Lakes, NJ) (14). The second sputum sample was stored (liquefied immediately or as raw sputum) at  $-20^{\circ}\text{C}$  for later analysis using the Xpert MTB/RIF assay. Patients who returned for postenrolment follow-up provided additional sputum samples at each visit. Smear grading according to the WHO/International Union Against Tuberculosis and Lung Disease method was performed (15). Culture-positive isolates underwent phenotypic drug susceptibility testing for rifampicin and isoniazid using the MGIT 960 SIRE kit (BD Diagnostics) (16). Cultures with a TTP of more than 28 days and only one out of four positive follow-up cultures (when available) were considered to be possible cross-contaminants (also analyzed separately) (7). Unless otherwise stated, all Xpert MTB/RIF and culture results were generated from paired samples taken at the same visit, and patients accordingly classified.

### Sample Processing and Storage for Later Analysis

The second sputa from the first 101 patients were liquefied using a 2:1 vol of 0.1% dithiothreitol (17) before storage at  $-20^{\circ}\text{C}$ . The remaining 395 samples were unprocessed and stored at  $-20^{\circ}\text{C}$  on collection.

### Sample Preparation and XpertMTB/RIF Procedure

Sputum sample preparation was performed as described previously (6, 18, 19) by a trained operator masked to clinical information. Briefly, the sample reagent (Cepheid) was mixed at a 2:1 ratio with 1 ml of sputum (either liquefied or unprocessed) and homogenized. If the sample volume was less than 1 ml, sterile phosphate buffer (Merck, Darmstadt, Germany) was added to bring the final volume to 1 ml. Two milliliters of homogenized mixture was transferred into an Xpert MTB/RIF assay cartridge and inserted into the GeneXpert instrument (6).

### Resolution of Discordant Results

For all Xpert MTB/RIF and culture discordant results, the cartridge-generated amplicon was extracted, amplified, and sequenced as previously described (6). In addition, a GenoTypeMTBDR<sub>plus</sub> test (HainLifescience, Nehren, Germany) or a PCR (using primers and conditions described previously) (6) followed by sequencing of the reaction products was performed on the stored sputum sediment or stored sample (18). Postenrolment sputum cultures were also analyzed and chest radiographs scored for likelihood of TB. Thus, performance was evaluated based on culture alone (Tables 1–4) or a combination of culture and these additional diagnostic investigations (Table 5). For the resolution of discordance in rifampicin resistance (phenotypic MGIT culture vs. Xpert MTB/RIF) a GenoTypeMTBDR-*plus* test was performed on the culture isolate. When appropriate, the *rpoB* gene from the sample sediment or the culture isolate was amplified and sequenced (6).

### Test Performance Assessment and Statistical Analysis

For the analysis of assay sensitivity, culture positivity and phenotypic susceptibility to rifampicin using simultaneously obtained paired samples (Xpert MTB/RIF vs. culture) were the reference standards. Specificity calculations were based on paired culture-negative samples from both culture-negative groups (probable and non-TB). Comparative specificity using the non-TB group only was also obtained. Test performance assessment and chi-square analyses were performed using OpenEpi (version 2.3.1; www.openepi.com) (20). Graphpad Prism (version 5.0; GraphPad Software, San Diego, CA) was used for the analysis of linear regression.

**TABLE 1. DEMOGRAPHIC INFORMATION AND CLINICAL CHARACTERISTICS STRATIFIED BY SMEAR STATUS**

Demographic or Clinical Characteristic	Study Cohort (%)	Smear-positive, Culture-positive (%)	Smear-negative, Culture-positive (%)	P Value
No. of TB suspects	480	94 (20)	47 (10)	
Median age (range)	36 (18–83)	35 (19–64)	37 (19–71)	0.56
Male	325 (68)	70 (22)	29 (9)	0.13
Female	155 (32)	24 (15)	18 (12)	0.13
Race				
Black	340 (71)	73 (21)	37 (12)	0.89
Mixed ancestry (colored)	132 (27)	19 (14)	10 (8)	0.88
White	9 (2)	2 (2)	0 (0)	0.44
Smoker (past or current)*	330 (73)	65 (20)	36 (11)	0.37
HIV positive†	130 (31)	23 (27)	23 (51)	<0.01
Median CD4 count (cells/ml) if HIV positive (range)‡	182 (0–935)	213 (0–439)	162 (10–465)	0.65
Previous TB§	158 (34)	24 (15)	16 (10)	0.30

Definition of abbreviation: TB = tuberculosis.

\* Excludes 26 patients with no smoking-related data.

† Excludes 59 patients who refused testing and 5 patients who had no data.

‡ Excludes seven patients who were HIV positive with no CD4 count data.

§ Excludes 19 patients with no data about previous TB history.

## RESULTS

### Patient Population

After the exclusion of samples from 16 patients (Figure 1), 480 patients with suspected TB were eligible for inclusion into the analysis. Figure 1 depicts how samples were processed and archived, and patients categorized by diagnostic subgroup. Patient demographic and clinical characteristics are shown in Table 1.

A total of 141 (29%) of 480 patients had definite TB. Of these, 94 (67%) of 141 patients had smear-positive, culture-positive TB, whereas 47 (33%) of 141 had smear-negative, culture-positive TB. A total of 182 (38%) patients were classified as probable

TB, whereas 157 (33%) patients were classified as non-TB. Of the 141 patients with positive cultures from their first sputum, 2 (1%) were found by phenotypic drug susceptibility testing to have MDR isolates (resistance to rifampicin and isoniazid), and 21 (15%) were found to have isoniazid mono-resistant isolates.

### Overall Xpert MTB/RIF Performance

The performance of Xpert MTB/RIF versus liquid culture performed on a simultaneously obtained paired spot sputum sample supplied at enrolment is shown in Table 2. The overall sensitivity of the assay was 78.7% (95% confidence interval [CI], 71.3–84.7;

**TABLE 2. PERFORMANCE OUTCOMES OF XPRT MTB/RIF FOR THE DETECTION OF MYCOBACTERIUM TUBERCULOSIS COMPARED TO SMEAR MICROSCOPY, AND STRATIFIED BY HIV STATUS**

	All Patients (n = 480)		Patients Uninfected with HIV (n = 286)*		Patients Infected with HIV (n = 130)*	
	Sens. (95% CI)†	Spec. (95% CI)‡	Sens. (95% CI)	Spec. (95% CI)‡	Sens. (95% CI)§	Spec. (95% CI)‡
Sputum smear	66.7 (58.5–73.9) 94 of 141	99.7 (98.4–100) 338 of 339	73.2 (62.7–81.6) 60 of 82	100 (98.2–100) 204 of 204	50 (36.1–63.9) 23 of 46 (P = 0.01)	98.8 (94.6–99.8) 83 of 84
Xpert MTB/RIF	78.7 (71.3–84.7) 111 of 141 (P = 0.02)	94.4 (91.4–96.4) 320 of 339	82.9 (73.4–89.6) 68 of 82	95.6 (91.8–97.7) 195 of 204	69.6 (55.2–80.1) 32 of 46 (P = 0.09)	91.7 (83.8–95.9) 77 of 84
Sputum smear or Xpert MTB/RIF	82.3 (75.1–87.7) 116 of 141 (P < 0.01)	94.1 (91.1–96.2) 319 of 339	85.4 (76.1–91.4) 70 of 82	95.6 (91.8–97.7) 195 of 204	73.9 (59.7–84.4) 34 of 46	87.5 (75.3–94.1) 76 of 84
Xpert MTB/RIF in smear-negative, culture-positive cases	46.8 (33.3–60.8) 22 of 47	N/A	45.5 (26.9–65.3) 10 of 22	N/A	47.3 (29.2–67) 11 of 23	N/A
	PPV (95% CI)	NPV (95% CI)†	PPV (95% CI)	NPV (95% CI)	PPV (95% CI)	NPV (95% CI)§
Sputum smear	99 (94.3–99.8) 94 of 95	87.8 (84.1–90.7) 338 of 385	100 (93.4–100) 60 of 60	90.3 (85.7–93.5) 204 of 226	95.8 (79.8–99.3) 23 of 24	78.3 (69.5–85.1) 83 of 106 (P < 0.01)
Xpert MTB/RIF	85.4 (78.3–90.4) 111 of 130	91.4 (88–93.9) 320 of 350	88.3 (79.3–93.4) 68 of 77	93.3 (89.1–96) 195 of 209	82.1 (67.3–91) 32 of 39	84.6 (75.8–90.6) 77 of 91 (P = 0.02)
Sputum smear or Xpert MTB/RIF	85.3 (78.4–90.3) 116 of 136	92.7 (89.5–95) 319 of 344 (P = 0.03)	88.6 (79.8–93.9) 70 of 79	94.2 (90.1–96.7) 195 of 207	85 (70.1–93) 34 of 42	77.8 (65.1–86) 76 of 88 (P < 0.001)
Xpert MTB/RIF in smear-negative, culture-positive cases	53.4 (38.8–68) 22 of 41	92.7 (89.5–95) 319 of 344	52.6 (31.7–72.3) 10 of 19	94.2 (90.1–96.7) 195 of 207	61.1 (38.6–79.7) 11 of 18	86.4 (77.7–92) 76 of 88 (P = 0.03)

Definition of abbreviations: CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value.

\* Excludes 59 patients who refused testing and 5 patients who had no data.

† P values < 0.10 in the "all patients" category are shown for comparisons between assays (microscopy vs. Xpert MTB-RIF, or microscopy vs. a combination of both).

‡ Specificity calculations were based on culture-negative samples obtained from both culture-negative groups (probable and non-TB)

§ Assay-specific (microscopy or Xpert MTB-RIF or a combination of both) P values < 0.10 comparing patients infected versus patients uninfected with HIV.

**TABLE 3. PERFORMANCE OUTCOMES OF XPERT MTB/RIF FOR THE DETECTION OF MYCOBACTERIUM TUBERCULOSIS COMPARED WITH SMEAR MICROSCOPY IN PERSONS INFECTED WITH HIV, AND STRATIFIED BY CD4 COUNT**

	Patients Infected with HIV (n = 130)*		Patients Infected with HIV with CD4 count ≥ 200 cells/ml† (n = 57)		Patients Infected with HIV with CD4 count <200 cells/ml (n = 66)	
	Sens. (95% CI)	Spec. (95% CI)‡	Sens. (95% CI)	Spec. (95% CI)‡	Sens. (95% CI)§	Spec. (95% CI)
Sputum smear	50 (36.1–63.9) 23 of 46	98.8 (94.6–99.8) 83 of 84	61.9 (40.1–79.3) 13 of 21	97.2 (85.8–99.5) 35 of 36	39.1 (22.2–59.2) 9 of 23 (P < 0.01)	100 (91.8–100) 43 of 43
Xpert MTB/RIF	69.6 (55.2–80.1) 32 of 46	91.7 (83.8–95.9) 77 of 84	76.2 (54.9–89.4) 16 of 21	97.2 (85.8–99.5) 35 of 36	65.2 (44.9–81.2) 15 of 23	93 (81.4–97.6) 40 of 43
Sputum smear or Xpert MTB/RIF	73.9 (59.7–84.4) 34 of 46	87.5 (75.3–94.1) 76 of 84	81 (60–92.3) 17 of 21	4.3 (81.4–98.4) 33 of 35	69.6 (49.3–84.4) 16 of 23	93 (81.4–97.6) 40 of 43
Xpert MTB/RIF in smear-negative, culture-positive cases	47.3 (29.2–67) 11 of 23	N/A	50 (21.5–78.5) 4 of 8	N/A	50 (26.8–73.2) 7 of 14	N/A
	PPV (95% CI)	NPV (95% CI)	PPV (95% CI)	NPV (95% CI)	PPV (95% CI)	NPV (95% CI)§
Sputum smear	95.8 (79.8–99.3) 23 of 24	78.3 (69.5–85.1) 83 of 106	92.9 (68.5–98.7) 13 of 14	81.4 (67.4–90.3) 35 of 43	100 (70–100) 9 of 9	75.4 (62.9–84.8) 43 of 57 (P < 0.01)
Xpert MTB/RIF	82.1 (67.3–91) 32 of 39	84.6 (75.8–90.6) 77 of 91	94.1 (73–99) 16 of 17	87.5 (73.9–94.5) 35 of 40	83.3 (60.8–94.2) 15 of 18	83.3 (70.4–91.3) 40 of 48 (P = 0.04)
Sputum smear or Xpert MTB/RIF	85 (70.1–93) 34 of 42	77.8 (65.1–86) 76 of 88	89.5 (68.8–97.1) 17 of 19	89.2 (75.3–95.7) 33 of 37	84.2 (71–96) 16 of 19	85.1 (72.3–92.6) 40 of 47 (P > 0.05)
Xpert MTB/RIF in smear-negative, culture-positive cases	61.1 (38.6–79.7) 11 of 18	86.4 (77.7–92) 76 of 88	80 (37.6–96.4) 4 of 5	89.5 (75.9–95.8) 34 of 38	70 (39.7–89.2) 7 of 10	85.1 (72.3–92.6) 40 of 47 (P > 0.05)

Definition of abbreviations: CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value

\* Excludes seven HIV-positive patients with no CD4 count data.

† Assay-specific performance in the patients infected with HIV with a CD4 count greater than 200 cells/ml group did not differ significantly compared with any other group.

‡ Specificity calculations were based on culture-negative samples obtained from both culture-negative groups (probable and non-TB).

§ Assay-specific P values (microscopy or Xpert or a combination of both) less than 0.10 comparing patients infected with HIV with a CD4 count less than 200 cells/ml versus patients uninfected with HIV.

111 of 141). In smear-positive culture-positive cases the sensitivity was 94.7% (88.2–97.7; 89 of 94), whereas in smear-negative culture-positive cases it was 46.8% (33.3–60.8; 22 of 47).

The overall specificity for the diagnosis of TB using both culture-negative groups (probable and non-TB) was 94% (91.4–96.4; 320 of 339). Using only the non-TB group, Xpert MTB/RIF specificity was 95% (91.7–98.4; 132 of 137; P = 0.39). The assay was negative in all eight cases that were culture-positive for non-TB mycobacteria (Figure 1), and negative in 14 of 15 cases with sputum isolates that were contaminated by bacterial overgrowth. Only 1 (0.2%) of 496 evaluated samples yielded an indeterminate Xpert MTB/RIF result.

A single Xpert MTB/RIF assay outperformed smear microscopy and showed an 18% relative increase in the rapid (potentially within 24 h) TB case-detection rate (17 additional cases) compared with 94 smear-positive cases and thus detected significantly more patients than smear microscopy (111 [78.7%] of 141 vs. 94 [66.7%] of 141; P = 0.02).

When a positive smear microscopy or Xpert MTB/RIF result were combined, the sensitivity improved further to 82.2% (75.1–87.7; 116 of 141) compared with smear microscopy alone (66.7% [68.5–73.4; 94 of 141]; P < 0.01) (Table 2).

#### Performance of Xpert MTB/RIF in Patients Infected with HIV

Smear microscopy was significantly less sensitive in subjects infected with HIV versus subjects uninfected with HIV (23 [50%] of 46 vs. 60 [73.2%] of 82; P = 0.01) (Table 2). Although the sensitivity of Xpert MTB/RIF was lower in the HIV-infected group, this did not reach significance (32 [69.6%] of 46 vs. 68 [82.9%] of 82; P = 0.09). The same pattern was seen in those with a CD4 count above or equal to versus below 200 cells/ml. Sensitivity of Xpert MTB/RIF in the smear-negative group was unaffected by HIV status or CD4 count

(Tables 2 and 3). By contrast, the negative predictive value (NPV) of Xpert MTB/RIF decreased significantly in patients infected with HIV versus patients uninfected with HIV (195 [93.3%] of 209 vs. 77 [84.6%] of 91; P = 0.02), and was lower in those with a CD4 count less than 200 cells/ml versus those with a CD4 count greater than or equal to 200 cells/ml (40 [83.3%] of 48 vs. 34 [87.5%] of 40; P = 0.60). The same pattern was seen for smear microscopy.

When the assays were directly compared within patient subgroups, the sensitivity of Xpert MTB/RIF was higher than smear microscopy in persons infected and uninfected with HIV, and in those with a CD4 count less than 200 cells/ml, but this difference was not significant for all three groups (Tables 2 and 3). The NPV for Xpert MTB/RIF did not differ significantly from that of smear microscopy in any of these groups.

The combination of smear microscopy and Xpert MTB/RIF had a significantly better sensitivity than smear microscopy alone in patients infected with HIV (34 [73.9%] of 46 vs. 23 [50%] of 46; P = 0.02) and in those with a CD4 count less than 200 cells/ml (16 [69.6%] of 23 vs. 9 [39.1%] of 23; P < 0.05). Likelihood ratios stratified by smear status, HIV status, and CD4 count are included in the online supplement.

#### Sample Processing and Volume, and Impact of Bacterial Burden

Sensitivity (77.9% vs. 81.1%; P = 0.70) and specificity (94.3% vs. 94.7%; P > 0.99) was similar in 386 unprocessed compared with 94 liquefied samples (Table 4). Similarly, sputum sample volume had limited impact on sensitivity (71.2% in samples less than the recommended 1 ml [median 575 µl; interquartile range {IQR} 300–700 µl] vs. 83.2% in samples of 1 ml [P = 0.10]; 40.9% in smear-negative culture-positive samples <1 ml [median 500 µl; IQR 300–600 µl] vs. 52% in samples of 1 ml [P = 0.47]). Restricting the analysis to include only unprocessed

**TABLE 4. IMPACT OF SAMPLE PROCESSING, SAMPLE VOLUME, AND CULTURE TIME TO POSITIVITY CUT-OFF ON XPERT MTB/RIF PERFORMANCE**

Sample Type		Sensitivity (95% CI)*			Specificity (95% CI) Number of Patients
		Culture-positive	Culture-positive, Smear-positive	Culture-positive, Smear-negative	
All	All samples (n = 480)	78.7 (71.3–84.7) 111 of 141	94.7 (88.2–97.7) 89 of 94	46.8 (33.3–60.8) 22 of 47	94.4 (91.4–96.4) 320 of 339
Sample processing	Unprocessed (n = 386)	77.9 (69–84.8) 81 of 104	93 (84.6–97) 66 of 71	45.5 (29.9–96) 15 of 33	94.3 (91–96.5) 266 of 282
	Liquefied (n = 94)	81.1 (65.8–90.5) 30 of 37	100 (85.7–100) 23 of 23	50 (26.8–73.2) 7 of 14	94.7 (85.6–98.2) 54 of 57
	P value <sup>†</sup>	0.70	0.24	0.78	0.95
	1 ml (n = 294)	83.2 (74–89.5) 74 of 89	95.3 (87.1–98.4) 61 of 64	52 (33.5–70) 13 of 25	95.1 (91.2–97.3) 195 of 205
Sample volume	<1 ml (n = 186) Median volume (IQR) = 575 µl (300–700)	71.2 (57.7–81.7) 37 of 52	93.3 (78.7–98.2) 28 of 30	40.9 (23.3–61.3) 9 of 22	93.3 (87.7–96.4) 125 of 134
	P value <sup>†</sup>	0.10	0.70	0.47	0.48
Culture time to positivity	Samples with a TTP >28 d excluded (n = 1) <sup>‡</sup>	79.3 (71.3–85.2) 111 of 140	94.7 (88.2–97.7) 89 of 94	47.8 (34.1–61.9) 22 of 46	94.4 (91.4–96.4) 320 of 339
All of the above filters	Unprocessed samples only with 1 ml of sputum available for the assay, and a culture TTP not >28 d with all subsequent sputum samples culture-negative (n = 241)	81.7 (71.2–90) 58 of 71	93.9 (83.5–97.9) 46 of 49	54.6 (34.7–73.1) 12 of 22	94.7 (71.2–90) 162 of 171
	P value <sup>†</sup>	0.62	0.83	0.56	0.89

Definition of abbreviations: CI = confidence interval; IQR = interquartile range; TTP = time-to-positivity.

\* Culture positivity for *M. tuberculosis* using a sputum sample obtained at enrolment was used as the reference standard.

<sup>†</sup> P values (two-tailed Mid-P chi-square test) comparing sputum processing method, sample volume, and samples with a TTP greater than 28 days excluded versus samples with a volume of 1 ml and a culture TTP not greater than 28 days (provided no subsequent cultures were positive).

<sup>‡</sup> Three patients had a first spot sputum sample culture TTP greater than 28 days. One patient gave a later sample that was culture-positive. One patient was lost to follow-up. For the remaining patients all three subsequent samples were culture-negative, indicating probable cross-contamination.

samples with 1 ml of available sputum and those with a culture TTP less than or equal to 28 days (provided no post-enrolment samples gave a positive culture) (21) did not significantly improve assay sensitivity (46.8%–54.6%;  $P = 0.56$ ) for smear-negative TB.

Higher bacterial loads, as determined by both smear grade and MGIT TTP, were associated with earlier detection by assay

and more frequent positive results (Figure 2). The average cycle threshold value was significantly lower in smear-positive compared with smear-negative cases ( $22 \pm 0.5$  vs.  $32 \pm 0.9$ ;  $P < 0.0001$ ). A similar relationship was seen using smear grade and culture TTP as markers of bacterial load (Figure 2). Using spiked sputum samples the limit of detection of the MTB/RIF assay was found to be 100 cfu/ml and freeze thaw

**TABLE 5. XPERT MTB/RIF PERFORMANCE OUTCOMES IN PATIENTS WITH DEFINITE TB (ALL CULTURE-POSITIVE CASES) VERSUS A GROUP CONTAINING CULTURE-POSITIVE AND HIGHLY LIKELY TB CASES\***

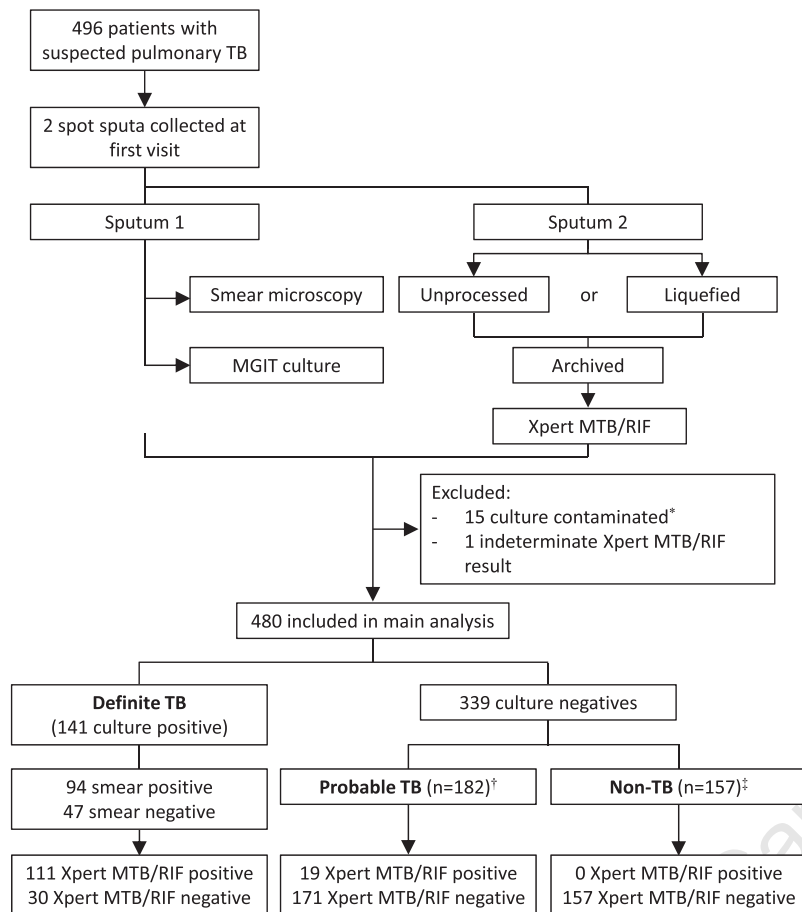
		Sens. (95% CI)	Spec. (95% CI)	PPV (95% CI)	NPV (95% CI)	Relative Increase in no. of Xpert MTB/RIF Diagnosed Cases Versus Smear Microscopy
All TB cases	Definite TB (culture-positive)	78.7 (71.3–84.7) 111 of 141	94.4 (91.4–96.4) 320 of 339	85.4 (78.3–90.4) 111 of 130	91.4 (88–93.9) 320 of 350	94–111 <sup>†</sup> (18%)
	Definite TB + highly likely group <sup>‡</sup>	80.9 (74–86.3) 127 of 157	99.1 (97.3–99.7) 320 of 323	97.7 (93.4–99.2) 127 of 130	91.4 (88–93.9) 320 of 350	94–126 (35%)
	P value	0.64	<0.001	<0.001	>0.99	<0.01
	Smear-negative TB	46.8 (33.3–60.8) 22 of 47 <sup>†</sup>	94.4 (91.4–96.4) 319 of 338	53.4 (38.8–68) 22 of 41	92.7 (89.5–95) 319 of 344	N/A
	Definite TB + highly likely group*	60.3 (48–71.5) 38 of 63	99.1 (97.3–99.7) 319 of 322	92.7 (80.6–97.5) 38 of 41	92.7 (89.5–95) 319 of 344	N/A
	P value	0.12	<0.001	<0.001	>0.99	N/A

Definition of abbreviations: CI = confidence interval; NPV = negative predictive value; PPV = positive predictive value; TB = tuberculosis.

\* P values indicate a comparison between groups. The same analysis for patients with smear-negative TB is also shown.

<sup>†</sup> Xpert MTB/RIF did not detect five smear-positive, culture-positive samples.

<sup>‡</sup> Sixteen culture-negative, Xpert MTB/RIF-positive patients are included here (five patients were found to be culture-positive by a second sputum obtained within 2 weeks from enrolment, five had *Mycobacterium tuberculosis* DNA in their sputum by sequencing, and six patients had typical radiologic evidence of active TB). All Xpert MTB/RIF amplicons were confirmed to contain *M. tuberculosis* DNA.



**Figure 1.** Flow diagram outlining patient enrolment, sample processing, and outcomes stratified by diagnostic category. TB = tuberculosis; MGIT= mycobacterial growth indicator tube.

\* 1 patient who had a contaminated spot sputum culture was Xpert<sup>®</sup> MTB/RIF-positive. No follow up culture results or radiological evidence was available for this patient.  
 † Includes 5 Xpert<sup>®</sup> MTB/RIF-negative individuals culture-positive for non-tuberculosis mycobacteria  
 ‡ Includes 3 individuals culture-positive for non-tuberculosis mycobacteria.

experiments showed no decrease in assay sensitivity (data not shown).

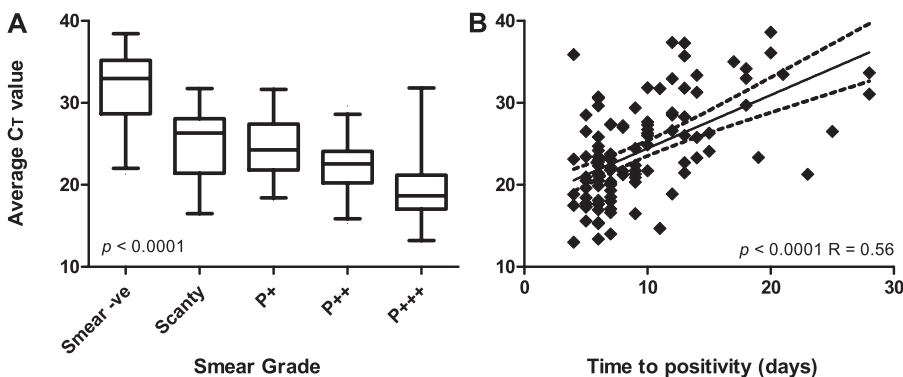
**Performance of Xpert MTB/RIF for the Detection of Rifampicin Resistance Using the Second-generation Software**

Xpert MTB/RIF identified six samples as rifampicin resistant. By contrast, MGIT DST identified five of these isolates as sensitive to rifampicin. Five out of the six samples were confirmed to be genotypically resistant by sequencing of DNA extracted from the

isolate or a GenoTypeMTBDR<sup>plus</sup> test. We were unable reliably to compute sensitivity given the small number of drug-resistant cases but Xpert MTB/RIF correctly determined susceptibility to rifampicin in 151 of 152 cases, and hence the specificity was 99.4% and the NPV was 98.7%.

**Discordance Between Xpert MTB/RIF and Culture (discordant analysis)**

*Xpert MTB/RIF-positive results in the probable group.* As shown in Figure 1, there were 19 Xpert MTB/RIF-positive patients



**Figure 2.** Correlation of the average cycle threshold (C<sub>T</sub>) value with smear grade (A) and liquid culture time to positivity (B).

who were culture-negative based on their simultaneously obtained paired sputum sample. Of these, five (26%) were found to be culture-positive on a second sputum obtained within 2 weeks of enrolment. A further 10 (53%) had *M. tuberculosis* DNA detected in their archived sputum using sequencing or a GenoTypeMTBDRplus test. In one (1%) additional patient the chest radiograph was compatible with and suggestive of active TB. Furthermore, in all 16 cases diagnosed with TB using either culture, sequencing of DNA from sputum or sample sediment, a GenoTypeMTBDRplus test, or chest radiography, the sequencing of the Xpert MTB/RIF cartridge amplicon confirmed *M. tuberculosis*. Thus, 16 of the 19 Xpert MTB/RIF-positive culture-negative patients were deemed likely to be true positives and designated as “highly likely TB.” In the remaining three patients a chest radiograph was unavailable precluding meaningful classification, although they were placed on treatment by the attending clinician based on clinical suspicion.

When the 16 of 19 patients were combined with the definite TB group and the data were reanalyzed (Table 5), there was no significant change in sensitivity (111 [78.7%] of 141 to 127 [80.9%] of 157;  $P = 0.64$ ) (Table 5). However, specificity (320 [94.4%] of 339 increased to 320 [99.1%] of 323;  $P < 0.001$ ) and the positive predictive value improved significantly (111 [85.4%] of 130 to 127 [97.7%] of 130;  $P < 0.001$ ), and the relative increase in the proportion of patients diagnosed compared with smear microscopy improved significantly (94–111 [18%] vs. 94–127 [35%];  $P < 0.01$ ). The positive predictive value in individuals with smear-negative TB improved significantly (22 [53.4%] of 41 to 38 [92.7%] of 41;  $P < 0.001$ ) and the sensitivity increased to 60.3%. Patients in the culture-negative Xpert MTB/RIF-positive patient group had a higher mean average cycle threshold value compared with those who were culture-positive, Xpert MTB/RIF-positive (29.3 [23.9–34] vs. 23.6 [19.6–27.3];  $P < 0.001$ ).

**Xpert MTB/RIF-negative culture-positive results.** Thirty patients were culture-positive but Xpert MTB/RIF-negative. There was a higher proportion of smear-negative individuals in this group compared with culture-positive, Xpert MTB/RIF-positive individuals (83% vs. 20%;  $P < 0.001$ ). The median TTP (IQR) was significantly longer in this group (18 [13–25] vs. 7 [6–12] d;  $P < 0.001$ ). There was no significant difference in median sample volume across these groups (0.9 vs. 1 ml;  $P = 0.13$ ).

## DISCUSSION

The WHO recently endorsed Xpert MTB/RIF (11); however, there are limited data about performance outcomes in high HIV prevalence settings where smear-negative TB is a formidable diagnostic challenge. The key findings of this preliminary study using archived samples were: (1) Xpert MTB/RIF outperformed smear microscopy because it diagnosed a significant proportion of smear-negative TB cases, and increased the relative proportion of potentially rapidly diagnosed cases by 18%; (2) HIV coinfection was associated with a significantly reduced assay NPV and there was a trend to reduced sensitivity; (3) taking into account Xpert MTB/RIF-positive culture-negative samples obtained from highly likely patients with TB the proportion of potentially rapidly diagnosed cases relative to smear microscopy significantly improved from 18–35%; (4) sputum volume and processing methods had a nonsignificant impact on assay performance, but by contrast bacterial load correlated strongly with performance; and (5) the specificity and NPV of the second-generation software for rifampicin resistance was almost 100%. Thus, Xpert MTB/RIF outperformed smear microscopy and simultaneously ruled out rifampicin resistance with great accuracy.

There are limited data about Xpert MTB/RIF performance in persons infected with HIV and none stratified by CD4 count. Our preliminary data indicate that the NPV is significantly reduced in this group (~15% of those with negative results have TB). The effects are most marked in those with advanced immunosuppression. Thus, in persons infected with HIV Xpert MTB/RIF is a good rule-in test but may have limited rule-out value compared with persons uninfected with HIV. This may be caused by the lower concentration of mycobacteria in the sputum of persons infected with HIV and possibly reduced specificity caused by occult or subclinical disease. Our data add to the limited existing knowledge base about the impact of HIV on Xpert MTB/RIF performance. However, there are several important limitations regarding our data (discussed in detail later) and thus further studies are required to clarify these findings. The added rule-out value of a second test in persons infected with HIV who are smear- and Xpert MTB/RIF-negative remains to be determined.

The key advantage of Xpert MTB/RIF is that it diagnosed 47% (95% CI, 33–61%) of smear-negative TB cases in a high HIV prevalence setting and 55% (35–73%) when a restricted analysis was performed. A recent multicenter study showed a sensitivity in the smear-negative group when using a single cartridge of 73% (65–79%) (7). The differing sensitivities likely reflect differences in study design or represent a chance finding given that the CI overlap in both the studies ( $P = 0.10$  vs. the restricted samples). Our results may reflect the effect of using frozen samples; however, preliminary experiments showed no significant effect of repeated freeze–thaw cycles on assay performance and similar observations have been reported by Helb and colleagues (6). That early morning sputum samples (known to have a higher diagnostic yield relative to spot sputum samples) (22, 23) were included in the study by Boehme and colleagues (7) and classification using at least one of four culture results per individual (which would decrease the number of culture-negative, Xpert MTB/RIF-positive cases) are other possible reasons for this discrepancy. Nevertheless, even with our detected sensitivity in the smear-negative group, there was an almost 20% increase in TB case detection compared with smear microscopy alone. This increase is dependent on local HIV prevalence rates.

Previous studies have not evaluated the significance of culture-negative Xpert MTB/RIF-positive samples. Are these true or false-positive results? When the significance of these results was clarified using short-term follow-up cultures, sequencing, and a suggestive radiologic picture using a standardized scoring system there was a 35% relative increase in the number of detected cases. Given the well-known limitations of *post hoc* discrepancy analyses (24), larger studies are now required in different settings to evaluate better discordant cases with long-term follow-up.

Our data suggest that sensitivity in smear-negative TB was limited by bacterial load. Studies correlating Xpert MTB/RIF Ct values with bacterial load are important because they inform contact tracing policies, treatment monitoring, and definition of a benchmarking threshold against which competitor assays can be measured. The limit of detection in our hands using spiked sputum samples was 100 CFU/ml. This is in keeping with the findings of Helb and colleagues (6), who report a limit of detection of 132 cfu/ml. Nevertheless, published studies confirm that those with smear-negative TB often have bacterial loads substantially below 100 cfu/ml (25). Future studies are required to examine if sample concentration can improve sensitivity.

A key advantage of Xpert MTB/RIF over smear microscopy is the simultaneous assessment for rifampicin resistance. We are unable to comment on sensitivity for rifampicin resistance given

the limited number of cases in this category, but we can confirm, similar to the findings of Boehme and colleagues (7) using the first-generation software, that the specificity and NPV are high (7). Thus, the assay using the second-generation software can reliably rule out rifampicin resistance in a high HIV prevalence setting. This is crucial given the increasing burden of MDR-TB and XDR-TB in Africa (26–28).

A major hurdle to widespread implementation of the Xpert MTB/RIF assay in resource-poor settings is cost (29). Thus, a possible interim strategy to enhance uptake might be to perform the assay only in smear-negative rather than smear-positive patients suspected of having TB (8, 28). Our data lend credence to this strategy because a combination of the two diagnostic methods showed the best sensitivity and specificity. However, the downside is that information about drug susceptibility is unavailable in smear-positive patients.

There are no existing published data on the effect of sputum volume on Xpert MTB/RIF performance. This is an important consideration in Africa, where HIV coinfection may result in a higher proportion of sputum-scarce patients who produce sub-optimal sputum volumes. Our data indicate that volumes below the recommended limit of 1 ml (median volume, 575  $\mu$ l) do have some impact on performance, although not significant, and thus could still be used for Xpert MTB/RIF in sputum-scarce patients. However, further studies are required to confirm these findings and to accurately determine the minimum volume of sputum that can be reliably used in sputum-scarce or paucibacillary individuals. It is possible that using volumes greater than 1 ml improves detection in smear-negative TB. That an approximately twofold reduction in volume has minimal impact on performance is not surprising given that PCR amplifies its DNA target by over a billionfold (30). Similarly, liquefied sputum had a minimal effect on performance. These data are important because in cases where the Xpert MTB/RIF assay is indeterminate or if an additional test is required then the decontaminated stored sample can be used to clarify the status of the donor.

There are several limitations of our study findings. First, our results may reflect the bias of sample storage and freeze–thaw, which may impact on DNA integrity or sputum viscosity. However, this is not supported by our preliminary freeze–thaw data or the data of Helb and colleagues (6), which suggest that prolonged storage and freeze–thaw have limited impact on sensitivity. Nevertheless, even under less rigorous and well-controlled study conditions compared with published data (7) we show that Xpert MTB/RIF comprehensively outperforms smear microscopy. We only performed one paired sputum-Xpert MTB/RIF culture per patient and used only a single liquid culture as a reference standard. Thus, we may have erroneously estimated outcomes given that a MGIT culture on a second specimen, compared with a MGIT culture on a single specimen from persons infected with HIV, may detect an additional 17% of cases (31). Our study findings may have limited relevance to low HIV prevalence settings where TB prevalence rates are lower (in patients infected with HIV and patients uninfected with HIV). The conclusiveness of our findings is especially limited by the small number of patients, particularly in the smear-negative and HIV-infected subgroups, which is a consequence of the study design. Prospective studies in these patient subgroups are now urgently needed. However, we establish a firm rationale and provide a foundation for the design of larger and more comprehensive studies to evaluate Xpert MTB/RIF in populations infected with HIV. Finally, our data do not inform on how Xpert MTB/RIF tests perform at the point-of-treatment, where the assay can have the greatest impact on patient care. Thus, controlled studies

evaluating outcomes at point-of-treatment in high HIV prevalent settings are urgently required.

In summary, the Xpert MTB/RIF assay is an accurate rapid rule-in test for pulmonary TB. It outperformed smear microscopy given that it established a diagnosis in a significant proportion of patients who are smear-negative. It may also detect additional culture-negative patients and has excellent rule-out value for MDR TB. However, because sample size, volume, and use of frozen samples were important limitations of this study, further studies are required to clarify test rule-out value in persons infected with HIV and the role of this assay in current TB treatment algorithms in high HIV prevalence settings.

**Author Disclosure:** G.T. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. J.P. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. R.V.Z.S. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. H.M.'s institution received a European Union funded FP7 grant for scientific training of student for capacity development. E.S. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. S.M. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. R.D. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. A.W. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. M.H.'s institution has received grants from the US Military HIV Research Program for HIV vaccine research, but also operational research on how to deliver care and treatment; from the European Commission for immunologic work on HIV/TB/Malaria research and helminth infection; from the European & Developing Countries Clinical Trials Partnership (EDCTP) for the conduct of clinical trials for TB and HIV drugs and vaccines, and the evaluation of new diagnostic tools for TB in adults and children; from the Federal Ministry of Education and Research that is cofounding to the previously mentioned EDCTP grants; from the Bill & Melinda Gates Foundation (BMGF) that is cofounding to some the previously mentioned EDCTP grants; and the World Health Organization for TB Reach funding to identify hard-to-reach patients with TB. S.S.'s institutions received a European Union funded FP7 grant for scientific training of student for capacity development. M.P. serves as a consultant to the BMGF. The BMGF had no involvement in this study. R.V. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript. K.D. does not have a financial relationship with a commercial entity that has an interest in the subject of this manuscript.

**Acknowledgment:** The authors are grateful to the patients who consented to take part in this study, Daliwonga Siganga and his team for overseeing patient recruitment, and Lwazi Mhlanti and Vonnita Louw for data capturing assistance. The authors are also grateful to Catharina Boehme and Pamela Nabeta of FIND for their helpful advice and technical assistance.

## References

1. WHO. Global tuberculosis control. Publication No. WHO/HTM/2010.7. Geneva, Switzerland: World Health Organization; 2010.
2. WHO. Multidrug and extensively drug-resistant TB (M/XDR-TB). Global report of surveillance and response. Publication No. WHO/HTM/TB/2010.3. Geneva, Switzerland: World Health Organization; 2009.
3. Getahun H, Harrington M, O'Brien R, Nunn P. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: informing urgent policy changes. *Lancet* 2007;369:2042–2049.
4. Pai M, Kalantri S, Dheda K. New tools and emerging technologies for the diagnosis of tuberculosis: part II. Active tuberculosis and drug resistance. *Expert Rev Mol Diagn* 2006;6:423–432.
5. New Diagnostic Working Group of the Stop TB Partnership. Pathways to better diagnostics for tuberculosis. A blueprint for the development of TB diagnostics. Geneva, Switzerland: World Health Organization; 2009.
6. Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, Kop J, Owens MR, Rodgers R, Banada P, *et al.* Rapid detection of *Mycobacterium tuberculosis* and rifampin resistance by use of on-demand, near-patient technology. *J Clin Microbiol* 2010;48:229–237.
7. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, *et al.* Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med* 2010; 363:1005–1015.
8. Van Rie A, Page-Shipp L, Scott L, Sanne I, Stevens W. Xpert() MTB/RIF for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype or hope? *Expert Rev Mol Diagn* 2010;10:937–946.

9. El-Hajj HH, Marras SAE, Tyagi S, Kramer FR, Alland D. Detection of rifampin resistance in *Mycobacterium tuberculosis* in a single tube with molecular beacons. *J Clin Microbiol* 2001;39:4131.
10. WHO. Roadmap for rolling out Xpert MTB/RIF for rapid diagnosis of TB and MDR-TB. Geneva, Switzerland: World Health Organization; 2010.
11. WHO. STAG-TB Report of the Tenth Meeting. Publication No. WHO/HTM/2010.18. Geneva, Switzerland: World Health Organization; 2010.
12. Dawson R, Masuka P, Edwards D, Bateman E, Bekker L, Wood R, Lawn S. Chest radiograph reading and recording system: evaluation for tuberculosis screening in patients with advanced HIV. *Int J Tuberc Lung Dis* 2010;14:52–58.
13. De Boon S, Bateman E, Enarson D, Borgdorff M, Verver S, Lombard C, Iruen E, Beyers N, White N. Development and evaluation of a new chest radiograph reading and recording system for epidemiological surveys of tuberculosis and lung disease. *Int J Tuberc Lung Dis* 2005;9:1088–1096.
14. Cruciani M, Scarparo C, Malena M. Meta-analysis of BACTEC MGIT 960 and BACTEC 460 TB, with or without solid media, for detection of mycobacteria. *J Clin Microbiol* 2004;42:2321–2325.
15. Van Rie A, Fitzgerald D, Kabuya G, Van Deun A, Tabala M, Jarret N, Behets F, Bahati E. Sputum smear microscopy: evaluation of impact of training, microscope distribution, and use of external quality assessment guidelines for resource-poor settings. *J Clin Microbiol* 2008;46:897.
16. Scarparo C, Ricordi P, Ruggiero G, Piccoli P. Evaluation of the fully automated BACTEC MGIT 960 system for testing susceptibility of *Mycobacterium tuberculosis* to pyrazinamide, streptomycin, isoniazid, rifampin, and ethambutol and comparison with the radiometric BACTEC 460TB method. *J Clin Microbiol* 2004;42:1109.
17. Cashmore T, Peter J, van Zyl-Smit R, Semple P, Maredza A, Meldau R, Zumla A, Nurse B, Dheda K, Hill P. Feasibility and diagnostic utility of antigen-specific interferon- responses for rapid immunodiagnosis of tuberculosis using induced sputum. *PLoS ONE* 2010;5:327–331.
18. Blakemore R, Story E, Helb D, Kop J, Banada P, Owens MR, Chakravorty S, Jones M, Alland D. Evaluation of the analytical performance of the Xpert MTB/RIF assay. *J Clin Microbiol* 2010;48:2495–2501.
19. Xpert MTB/RIF [package insert]. Cepheid S, CA, 300–7810 Rev. A, April 2009.
20. Dean AGSK, Soe MM. OpenEpi: open source epidemiologic statistics for public health, Version 2.3.1. Available at: [www.OpenEpi.com](http://www.OpenEpi.com), updated 2010/19/09. Accessed November 30, 2010.
21. Rishi S, Sinha P, Malhotra B, Pal N. A comparative study for the detection of mycobacteria by BACTEC MGIT 960, Lowenstein Jensen media and direct AFB smear examination. *Indian J Med Microbiol* 2010;25:383.
22. Schoch O, Rieder P, Tueller C, Altpeter E, Zellweger J, Rieder H, Krause M, Thurnheer R. Diagnostic yield of sputum, induced sputum, and bronchoscopy after radiologic tuberculosis screening. *Am J Respir Crit Care Med* 2007;175:80.
23. Mase S, Ramsay A, Ng V, Henry M, Hopewell P, Cunningham J, Urbanczik R, Perkins M, Aziz M, Pai M. Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: a systematic review. *Int J Tuberc Lung Dis* 2007;11:485–495.
24. Hadgu A. The discrepancy in discrepant analysis. *Lancet* 1996;348:592–593.
25. Chakravorty S, Tyagi J. Novel multipurpose methodology for detection of mycobacteria in pulmonary and extrapulmonary specimens by smear microscopy, culture, and PCR. *J Clin Microbiol* 2005;43:2697.
26. Dheda K, Shean K, Zumla A, Badri M, Streicher E, Page-Shipp L, Willcox P, John M, Reubenson G, Govindasamy D. Early treatment outcomes and HIV status of patients with extensively drug-resistant tuberculosis in South Africa: a retrospective cohort study. *Lancet* 2010;375:1798–1807.
27. Gandhi N, Nunn P, Dheda K, Schaaf H, Zignol M, van Soolingen D, Jensen P, Bayona J. Multidrug-resistant and extensively drug-resistant tuberculosis: a threat to global control of tuberculosis. *Lancet* 2010;375:1830–1843.
28. Dheda K, Warren R, Zumla A, Grobusch M. Extensively drug-resistant tuberculosis: epidemiology and management challenges. *Infect Dis Clin North Am* 2010;24:705–725.
29. Small PM, Pai M. Tuberculosis diagnosis: time for a game change. *N Eng J Med* 2010;363:1070–1071.
30. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation: strategies and considerations. *Genes Immun* 2005;6:279–284.
31. Monkongdee P, McCarthy KD, Cain KP, Tasaneeyapan T, Dung NH, Lan NTN, Yen NTB, Teeratakulpisarn N, Udomsantisuk N, Heilig C. Yield of acid-fast smear and mycobacterial culture for tuberculosis diagnosis in people with human immunodeficiency virus. *Am J Respir Crit Care Med* 2009;180:903.

## **Chapter 6.**

### **Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients**

Peter, J. G., Theron, G., van Zyl-Smit, R.N., Haripersad, A., Mottay, L., Kraus, S., Binder, A., Meldau, R., Hardy, A. and Dheda, K.

Published: *European Respiratory Journal* 2012 Nov; 40(5):1211-20

And

### **Urine antigen test for diagnosis of HIV-associated tuberculosis**

Peter, J.G., Theron, G. and Dheda, K.

Published: *Lancet Infectious Diseases* 2012 Nov; 12(11):825

### **PhD Context**

Despite improved sputum sampling strategies such as SI, up to a fifth of adults with suspected SN- or SSTB remain unable to produce a sputum sample. Furthermore, in high HIV prevalence settings extrapulmonary TB or disseminated TB account for up to a third of the TB cases. South African hospitals are overburdened with HIV-infected patients with advanced immunosuppression with suspected SN-, disseminated or extrapulmonary TB, and this patient group carries the highest TB-related mortality. The use of rapid, non-sputum based diagnostic tools, such as the urine LAM strip test, is essential to target this large vulnerable group. Thus, we undertook an evaluation of the

urine LAM strip (and ELISA) tests as an alternative diagnostic strategy complementary to improved sputum sample acquisition. This study of the urine LAM strip test, the first immunochromatographic point-of-care strip test developed for TB diagnosis, was the first evaluation of this assay in hospitalised patients with suspected TB HIV co-infection and advanced immunosuppression. Together with an evaluation in an outpatient ARV clinic setting, they represent the only two published evaluations of this assay to date. This study, with its evaluation and optimisation of test characteristics, represents an important advance in the progress of this test into clinical practice.

University of Cape Town



# Diagnostic accuracy of a urine lipoarabinomannan strip-test for TB detection in HIV-infected hospitalised patients

Jonathan G. Peter\*, Grant Theron\*, Richard van Zyl-Smit\*,  
Asheen Haripersad\*, Lynelle Mottay\*, Sarah Kraus\*, Anke Binder\*,  
Richard Meldau\*, Anneli Hardy# and Keertan Dheda\*,<sup>†,‡,+</sup>

**ABSTRACT:** Lack of point-of-care tests for tuberculosis (TB) result in diagnostic delay, and increased mortality and healthcare-related costs.

The urine Determine™ TB-LAM point-of-care strip-test was evaluated in 335 prospectively-recruited hospitalised patients with suspected TB-HIV co-infection (group 1) and from 88 HIV-infected hospitalised patients with non-TB diagnoses (group 2). Cut-off point-specific analyses were performed using: 1) a microbiological reference standard (culture positive versus negative); and 2) a composite reference standard (exclusion of patients with clinical-TB from the culture-negative group).

Using the microbiological reference and the manufacturer-recommended grade-1 cut-off point, LAM sensitivity and specificity was 66% (95% CI 57–74%). By contrast, using the composite reference sensitivity was 60% (95% CI 53–67%) and specificity improved to 96% (95% CI 89–100%) ( $p=0.001$ ). The same pattern was seen when the grade-2 cut-off point was used (specificity 75% versus 96%;  $p=0.01$ ). In group two patients specificity was poor using the grade-1 cut-off point, but improved significantly when the grade-2 cut-off point was used (90% versus 99%;  $p=0.009$ ). The grade-2 cut-off point also offered superior inter-reader reliability ( $p=0.002$ ). Sensitivity was highest in those with a CD4 <200 cells per mL. LAM combined with smear-microscopy was able to rule-in TB in 71% of *Mycobacterium tuberculosis* culture-positive patients.

This preliminary study indicates that the LAM strip-test may be a potentially useful rapid rule-in test for TB in hospitalised patients with advanced immunosuppression. The grade 2, but not the manufacturer-recommended grade 1 cut-off point, offered superior rule-in utility and inter-reader reliability. Larger studies to evaluate cut-off points and diagnostic accuracy are urgently required.

**KEYWORDS:** Diagnostic, lateral flow assay, lipoarabinomannan, point-of-care

The co-epidemics of tuberculosis (TB) and HIV remain out of control in sub-Saharan Africa [1]. TB is still the leading cause of HIV-related mortality [2] and up to 80% of newly diagnosed TB cases in sub-Saharan Africa are HIV co-infected [1]. African hospitals are overwhelmed by admissions of TB-HIV co-infected patients, hospital stay is prolonged because of suboptimal diagnostic tools and, consequently, healthcare-related costs are high [3]. A recent South African study showed that the early commencement of anti-TB treatment decreased 8-week mortality and length of stay in hospitalised TB patients [4].

An autopsy study of 240 in-patients (HIV seroprevalence rate of 94%) found that, of those not on anti-TB treatment, >40% were culture positive for *Mycobacterium tuberculosis* at the time of death [5]. Thus, delay or failure to diagnose TB is associated with high mortality in HIV co-infected hospitalised patients. Traditional TB diagnostic tools perform poorly in these patients [6, 7], and novel molecular diagnostics such as Xpert MTB/RIF (Cepheid, Sunnyvale, CA, USA) and other diagnostic adjuncts such as bronchoscopy, remain unaffordable and/or limited by the need for specialised laboratories and training [8]. Moreover, current tools perform

## AFFILIATIONS

\*Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Dept of Medicine, University of Cape Town,  
#Dept of Statistical Sciences, University of Cape Town, and  
†Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa.  
‡Dept of Infection, University College London Medical School, London, UK.

## CORRESPONDENCE

K. Dheda  
University of Cape Town  
H47 Old Main Building  
Groote Schuur Hospital  
Observatory  
7925 Cape Town  
Cape Town  
South Africa  
E-mail: Keertan.Dheda@uct.ac.za

Received:  
Nov 18 2011  
Accepted after revision:  
Feb 08 2012  
First published online:  
Feb 23 2012

This article has supplementary material available from [www.erj.ersjournals.com](http://www.erj.ersjournals.com)

poorly in sputum-scarce and extrapulmonary (EP) TB. There remains a pressing need for a rapid, inexpensive and simple point-of-care (POC) TB diagnostic for HIV co-infected patients.

Lipoarabinomannan (LAM) is a 17.5 kDa glycolipid that forms a component of the outer cell wall of mycobacterial species [9]. It is filtered by the kidney and detectable in the urine, and has been extensively evaluated as a TB diagnostic antigen [10–15]. A urine ELISA-based kit to detect LAM antigen (Clearview® TB-ELISA; Alere, Waltham, MA, USA) was found to be clinically useful in hospitalised TB-HIV co-infected patients with advanced immunosuppression [16], and was shown to be more sensitive than smear-microscopy in a TB screening study prior to the introduction of highly active antiretroviral therapy (HAART) [17].

The ELISA version of the test has now evolved into a simple POC urine immunochromatographic strip-test (lateral flow assay). The commercial version of the strip-test, which was evaluated in this study, is planned for release in the first quarter of 2013. This urine LAM strip-test is marketed by Alere, USA as the Determine® TB-LAM Ag rapid test (fig. 1). This novel test requires only 60 µL of unprocessed urine, can be performed at the bedside, requires little technical skill, no power source and provides a result within 25 min at a probable cost of less than US\$3.50 (the locked-in final version evaluated here is to be launched in quarter 1 of 2013) [18]. A recent evaluation of the LAM strip-test in outpatients being screened for antiretroviral therapy at primary care clinics showed promising results, especially for patients with advanced immunosuppression [19]. However, performance amongst hospitalised HIV-infected patients with advanced immunosuppression, in whom rapid diagnosis and treatment could potentially impact mortality remains unknown. Thus, we undertook a study to

prospectively evaluate the diagnostic accuracy of the LAM strip-test in hospitalised HIV co-infected patients with suspected TB. Some of the preliminary results of these studies, using the manufacturer-recommended cut-off point, have been previously reported in the form of an abstract [16].

## METHODS

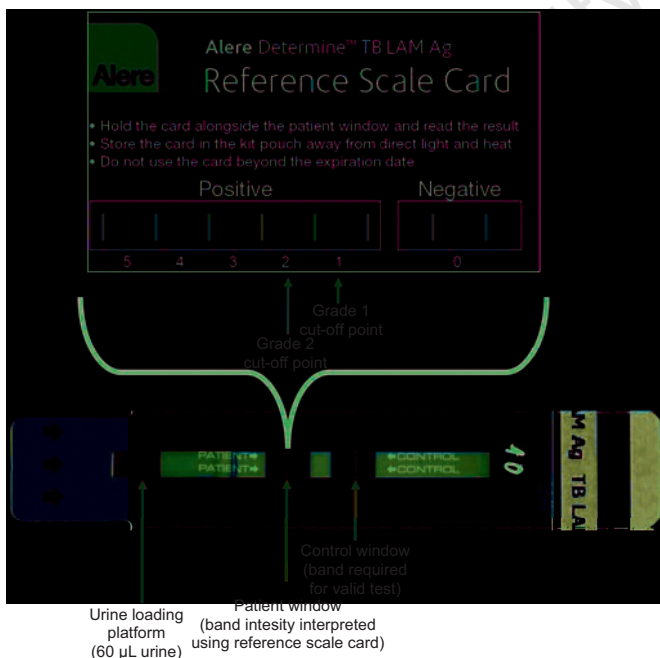
### Study population

The study population, divided into two groups, consisted of 423 adult patients recruited from four hospitals (three district- and one tertiary-level hospital; Groote Schuur, GF Jooste, Victoria and Somerset Hospitals, Cape Town, South Africa) between July 22, 2009 and December 14, 2010 in Cape Town, South Africa. Study group 1 consisted of 335 prospectively recruited hospitalised individuals with suspected TB and HIV, while study group 2 consisted of 88 HIV-infected hospitalised non-TB patients (possessing alternative diagnoses and no clinical or microbiological evidence of TB). Group 2 patients were recruited to further evaluate test specificity due to the limited number of confirmed non-TB patients in group 1. A study and analysis outline is shown in figure 2. Group 1 hospitalised patients were consecutively recruited following referral by clinicians who considered patients to be HIV-infected and with suspected TB. Group 2 patients represented a convenience sample and were specifically recruited as they were not suspected of having TB. Only three patients, who refused consent, were excluded from the study. All patients provided informed consent and the study was approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee. Clinical information documented for enrolled patients included demographic information, past history of TB, symptoms and vital signs, HIV status, renal function, and details about anti-TB therapy.

The clinical management of all patients (including the timing and extent of diagnostic work-up, decision to commence anti-TB treatment and final discharge from hospital) was undertaken by consultant-led groups of health care workers who had no association with the study team. TB diagnostic work-up was not standardised and was at the discretion of the attending senior physician. Table 1 provides a site-specific breakdown of clinical samples collected for TB diagnosis. The local reference laboratory processed all clinical specimens collected for routine TB diagnosis (fluorescence smear microscopy on NALC/NaOH processed and concentrated samples followed by liquid culture (BACTEC MGIT 960; BD Diagnostics, Franklin Lakes, NJ, USA) were performed when appropriate).

### LAM methodology

All patients were required to give a spot urine sample (10–30 mL) collected in a sterile container as soon as possible after recruitment. A urine dipstick test (UriCHECK 9; RapiMed Diagnostics, Pretoria, South Africa) was immediately performed to assess for protein, blood and leukocytes. Urine was stored at -20°C for later batched testing. The LAM strip-test (a single manufacturing lot #101102) was performed on unprocessed thawed urine according to manufacturer's instructions and after 25–35 min, two readers, blinded to clinical details and TB status, independently evaluated the LAM strips for all study patients. The detailed methods of LAM strip-test grading using the manufacturer's reference card (fig. 1), the interpretation of



**FIGURE 1.** Lipoarabinomannan (LAM) strip-test and reference scale card. The reference scale card (courtesy of Alere, Waltham, MA, USA) illustrates six cut-off points (visual grades 0–5) categorised by different band intensities appearing in the patient window.



**TABLE 1** Demographic, clinical and microbiological characteristics of HIV-infected patients in the study population stratified by *Mycobacterium tuberculosis* culture status, CD4 cell count and patient group

	All	Group 1 patients			Group 2 patients		p-value†
		<i>M. tuberculosis</i> culture positive	<i>M. tuberculosis</i> culture negative	CD4 ≤200 cells per mL‡	CD4 >200 cells per mL‡	Non-TB (control)	
<b>Subjects n</b>	242	116	126	173	55	88	
<b>Demographics</b>							
Age yrs	35 (29–40)	35 (29–39)	35 (30–40)	35 (30–40)	34 (29–38)	36 (30–42)	NS
Male	96 (40)	47 (41)	49 (39)	72 (42)	15 (27)	28 (32)	NS
CD4 count	90 (47–197)	86 (42–192)	91 (51–197) <sup>§</sup>	65 (35–112)	235 (235–408)	238 (140–404) <sup>§</sup>	<0.001 <sup>§</sup>
Previous TB	85(35)	33 (29) <sup>§</sup>	52 (41) <sup>§</sup>	58 (34)	18 (33)	32 (36)	0.04 <sup>§</sup>
Current smoker	47(19)	25 (22)	22 (18)	29 (17)	11(20)	11(13)	NS
<b>Clinical features</b>							
Cough >2 weeks	197 (81)	99 (85)	98 (78)	140 (81)	44 (80)	0 (0)	NS
Night sweats	163 (67)	82 (71)	81 (64)	116 (67)	36 (65)	0 (0)	NS
Weight loss	213 (88)	105 (91)	108 (86)	155 (90)	44 (80)	0 (0)	NS
Fever >38°C	44 (18)	29 (25) <sup>§</sup>	15 (12) <sup>§</sup>	33 (19)	8 (15)	0 (0)	0.008 <sup>§</sup>
<b>Clinical samples collected for TB culture</b>							
1 sputum sample	195 (81)	95 (82)	100 (79)	137 (79)	48 (87)	No results	NS
≥2 sputum samples	88 (36)	42 (36)	46 (37)	63 (42)	22 (40)	No results	NS
1 non-sputum sample	151 (62)	78 (67)	73 (58)	113 (65)	30 (55)	No results	NS
≥2 non-sputum sample	56 (23)	26 (22)	30 (24)	45 (26)	10 (18)	No results	NS
<b>Microbiological TB diagnosis</b>							
1 × sputum <i>M. tuberculosis</i> culture positive	89 (37)	89 (77)	0 (0)	59 (34)	24 (34)	No results	NS
1 × non-sputum culture <i>M. tuberculosis</i> positive*	47 (19) <sup>§</sup>	47 (41)	0 (0)	38 (22) <sup>§</sup>	5 (9) <sup>§</sup>	No results	0.03
1 × both sputum and non-sputum <i>M. tuberculosis</i> culture positive	20 (8)	20 (17)	0 (0)	16 (9)	3 (6)	No results	NS
Any sputum smear AFB-positive	65 (27)	60 (52)	5 (4)	40 (23)	15 (29)	No results	NS
Sputum scarce/smear-negative, <i>M. tuberculosis</i> culture positive	56 (23)	56 (48)	0 (0)	41 (23)	11 (20)	No results	NS
Non-sputum sample smear AFB-positive	15 (6)	8 (7)	7 (6)	11 (6)	2 (4)	No results	NS

Data are presented as median (interquartile range) or n (%), unless otherwise stated. TB: tuberculosis; AFB: acid-fast bacillus; ns: nonsignificant. †: excludes 14 HIV-positive patients with no CD4 cell count data; ‡: p-values indicate significant differences between patient groups (marked with § to indicate comparison group) for demographic, clinical or microbiological characteristics; †: includes the following *M. tuberculosis* culture-positive samples: seven blood cultures, 12 pleural fluid samples, four pericardial fluid samples, three ascitic fluid samples, eight cerebrospinal fluid samples, one hip biopsy, two lymph node biopsies, six fine-needle aspirates, one gastric washing, one faeces and two pus swabs.

**TABLE 2** Performance outcomes of smear microscopy, lipoarabinomannan (LAM) ELISA and LAM strip-test (using grade 1 and 2 cut-off points) in HIV-infected patients

Diagnostic test(s)	Group 1 patients				Group 2 non-TB control patients
	Microbiological reference ( <i>Mycobacterium tuberculosis</i> culture positive versus negative)		Composite reference standard (definite- and probable-TB versus non-TB)		Specificity
	Sensitivity	Specificity	Sensitivity	Specificity	
<b>Subjects n</b>	242		214		88
<b>Rapid smear microscopy</b>	56 <sup>#,†</sup> (47–65) 65/116	90 <sup>†</sup> (84–95) 114/126	44 <sup>#,***</sup> (37–51) 82/187	96 (89–100) 26/27	No results
<b>LAM ELISA</b>	59 (50–68) 68/115	80 <sup>*,†,‡,***</sup> (72–86) 98/123	51 (44–58) 93/183	96 <sup>§</sup> (89–100) 24/25	99 <sup>+</sup> (94–100) 87/88
<b>LAM strip-test (grade 1 cut-off point)</b>	66 <sup>**</sup> (57–74) 77/116	66 <sup>*,§,***</sup> (57–73) 82/125	60 (53–67) 112/187	96 <sup>##</sup> (89–100) 26/27	90 <sup>+‡</sup> (82–95) 79/88
<b>LAM strip-test (grade 2 cut-off point)</b>	50 <sup>***,***</sup> (41–59) 58/116	75 <sup>†,##,†</sup> (67–82) 94/125	45 <sup>***</sup> (39–53) 85/187	96 <sup>##</sup> (89–100) 26/27	99 <sup>+‡</sup> (94–100) 87/88
<b>LAM strip-test (grade 2 cut-off point) in smear-negative/sputum scarce <i>M. tuberculosis</i> culture-positive patients</b>	39 (28–52) 22/56	N/A	38 (29–48) 36/95	N/A	N/A
<b>Combined sputum smear microscopy and LAM strip-test (grade 2 cut-off point)</b>	71 <sup>†,***</sup> (62–78) 82/116	73 <sup>†,††</sup> (65–80) 92/126	63 <sup>***</sup> (56–70) 118/187	93 <sup>††</sup> (83–100) 25/27	99 <sup>+</sup> (94–100) 87/88

Data are presented as sensitivity or specificity (95% CI) n/n<sub>total</sub>, unless otherwise stated. N/A: not applicable. Using *M. tuberculosis* culture as the reference standard, 242 HIV-infected group 1 patients had  $\geq 1$  valid tuberculosis (TB) culture result and were included in this analysis. 116 and 126 *M. tuberculosis* culture-positive and -negative patients were used for sensitivity and specificity calculations, respectively. Using the composite reference standard, 67 out of 281 patients were TB unclassifiable and, hence were excluded from this analysis. 116 definite-TB combined with 71 probable-TB were used for sensitivity calculations and 27 non-TB were used for specificity calculations. Additionally, 88 group 2 non-TB control patients were used to calculate a second specificity. Rapid smear microscopy includes both sputum and non-sputum samples (e.g. fine-needle aspirate and cerebrospinal fluid samples) attainable within 24 h of hospital admission. One and three *M. tuberculosis* culture positive (analysis 1) and four definite-TB and two non-TB (analysis 2) had a missing LAM ELISA result. One *M. tuberculosis* culture negative patient had a missing urine LAM strip-test. For a comparison of the specific measures of diagnostic accuracy (sensitivity or specificity) between different tests (LAM ELISA versus LAM strip-test) or combinations thereof, or between different cut-off points of the urine LAM strip-test (grade 1 versus grade 2 cut-off point) p-values were  $p < 0.05$ ; specific p-values: <sup>†</sup>:  $p = 0.02$ ; <sup>\*\*</sup>:  $p = 0.01$ ; <sup>\*\*\*</sup>:  $p = 0.001$ ; <sup>‡</sup>:  $p = 0.009$ ; nonsignificant p-values not shown. For a comparison of differences in diagnostic accuracy measures between analyses and patient groups for a specific test (LAM ELISA or LAM strip-test) p-values were  $p < 0.05$ ; specific p-values: <sup>\*</sup>:  $p = 0.05$ ; <sup>##</sup>:  $p = 0.01$ ; <sup>†</sup>:  $p < 0.001$ ; <sup>‡</sup>:  $p = 0.04$ ; <sup>§</sup>:  $p = 0.001$ ; <sup>††</sup>:  $p = 0.03$ .

were calculated with 95% confidence intervals (positive likelihood ratios are presented in the online supplementary results). Intra- and interobserver agreement was assessed using the Kappa statistic, and categorised as poor (0–0.20), fair (0.21–0.40), moderate (0.41–0.60), good (0.61–0.80) and very good (0.81–1) [21]. Additionally, a McNemar Chi-squared test was used to compare the difference in agreement between readers at the grade 1 and 2 cut-off points. STATA IC, version 10 (StataCorp, College Station, TX, USA) was used for all statistical analyses. Study reporting and analysis were consistent with the Standards for the Reporting of Diagnostic Accuracy (STARD) criteria [22].

## RESULTS

### Demographic, clinical and microbiological characteristics of study patients (groups 1 and 2)

HIV infection was confirmed in 281 (84%) out of 335 of group 1 patients and consequently, 54 HIV-uninfected group 1 patients were excluded from the analyses. The demographic, clinical, diagnostic work-up and microbiological details of the HIV-infected

patients stratified by culture status, patient group and CD4 cell count are shown in table 1. The median (interquartile) CD4 cell count in group 1 *M. tuberculosis* culture-negative HIV-infected patients was significantly lower than the group 2 non-TB patients (91 (51–197) versus 238 (140–404) cells per mL;  $p < 0.001$ ). Of all clinical symptoms and signs, only an admission fever  $> 38^{\circ}\text{C}$  was associated with *M. tuberculosis* culture positivity ( $p = 0.008$ ).

116 (48%) out of 242 of group 1 HIV-infected patients were *M. tuberculosis* culture positive. Using the composite reference, a further 71 patients were categorised as probable-TB (nine sputum/non-sputum sample smear positive, *M. tuberculosis* culture negative/contaminated). Given suspected TB–HIV coinfection, attending clinicians attempted, where possible, to collect both sputum and non-sputum samples for TB diagnosis, with 195 (81%) out of 242 and 151 (62%) out of 242 of all group-1 HIV-infected patients having at least one sputum and non-sputum sample. Of the HIV-infected *M. tuberculosis* culture-positive patients, 69 (59%) out of 116 were sputum culture

**TABLE 3** Sensitivity of smear microscopy, lipoarabinomannan (LAM) ELISA and LAM strip-test (using grade 1 and 2 cut-off points) in different forms of *Mycobacterium tuberculosis* culture-positive disease<sup>#</sup>

Diagnostic test(s)	<i>M. tuberculosis</i> culture-positive group 1 patients		
	PTB only	PTB and EPTB	EPTB only
Subjects n	69	20	27
Rapid smear microscopy <sup>†</sup>	67*** (55–77) 46/69	70 (48–85) 14/20	19***,5,†,## (8–37) 5/27
LAM ELISA <sup>+</sup>	59 (47–70) 40/68	75 (53–89) 15/20	52 <sup>‡</sup> (34–69) 14/27
LAM strip-test (grade 1 cut-off point)	67 (55–77) 46/69	80 (58–92) 16/20	59 (41–76) 16/27
LAM strip-test (grade 2 cut-off point)	51 (39–62) 35/69	60 (39–78) 12/20	41 <sup>##</sup> (25–59) 11/27
LAM strip-test (grade 2 cut-off point) in smear-negative/ sputum scarce <i>M. tuberculosis</i> culture-positive patients	35 (19–55) 8/23	50 (19–81) 3/6	32 (16–53) 7/22
Combined sputum smear microscopy and LAM strip-test (grade 2 cut-off point)	78 (67–86) 54/69	85 (64–95) 17/20	44 <sup>†</sup> (28–63) 12/27

Data are presented as sensitivity (95% CI) n/n<sub>total</sub>, unless otherwise stated. PTB: pulmonary tuberculosis (TB); EPTB: extrapulmonary TB. <sup>#</sup>: 116 definite-TB patients were used for calculations (69 patients with at least one culture-positive sputum, 20 patients with at least one culture-positive sample from sputum and another non-sputum site and 27 culture positive from a non-pulmonary site, *i.e.* not from sputum). <sup>†</sup>: Rapid (within 24 h of hospitalisation) smear microscopy includes both sputum and non-sputum samples (*e.g.* fine-needle aspirate and cerebrospinal fluid samples). <sup>+</sup>: one *M. tuberculosis* culture-positive patient was missing a LAM ELISA result. For a comparison of test sensitivity in specific TB disease states between different tests (LAM ELISA *versus* LAM strip-test) or combinations thereof, or between different cut-off points of the urine LAM strip-test (grade 1 *versus* grade 2 cut-off point) p-values were <sup>‡</sup>: p=0.01; <sup>†</sup>: p=0.04; <sup>##</sup>: p=0.07 (if analysis restricted to sputum smear microscopy only then p<0.001). <sup>\*\*\*</sup>: p<0.001 for a comparison of a specific test (LAM ELISA or LAM strip-test) in different TB disease states.

positive only (PTB), 27 (23%) out of 116 were non-sputum culture positive only (EPTB) and 20 (18%) out of 116 were culture positive on both sputum and non-sputum samples (PTB/EPTB). Overall, 56 (48%) out of 116 of *M. tuberculosis* culture-positive patients were either sputum scarce or sputum smear-negative (21 (38%) out of 56 sputum scarce and 35 (62%) out of 56 sputum smear-negative).

The performance outcomes for rapid smear microscopy stratified by patient group, analysis 1 and 2, TB diagnostic subtype and CD4 cell count are shown in tables 2–4. Using the microbiological reference (analysis 1), the sensitivity (95% CI) of rapid smear microscopy was 56% (47–65%), while using the composite reference (analysis 2) sensitivity (95% CI) was 44% (37–51%) (p=0.04). The sensitivity (95% CI) of rapid smear microscopy was significantly lower in EPTB compared with PTB (19% (8–37%) *versus* 67% (55–77%), p<0.001) (table 3). Given the low diagnostic yield of smear microscopy, 118 (49%) out of 242 of unselected HIV-infected group 1 patients underwent specialised radiology to aid diagnostic decision-making (92 had abdominal ultrasound, 12 had computed tomography of the brain and 27 had echocardiography with/without pericardiocentesis).

#### Performance of the LAM strip-test using different cut-off points

Reader-specific performance outcomes and ROC analysis for both readers for analysis 1 and 2 are displayed in figure 3. In the analysis 1 (fig. 3a), using the manufacturer-suggested grade 1 cut-off point, the sensitivity, specificity and positive likelihood ratio (LR+) for reader 1 was 63.8%, 72.8% and 2.3, respectively, and for reader 2 was 66.4%, 65.6% and 1.9, respectively. Using a grade 2 or 3 cut-off point to optimise test rule-in value results in a reduced test sensitivity at both cut-off points (reader 1: 74 (63.8%) out of 116 *versus* 53 (45.7%) out of 116 (grade 2) and 47 (40.5%) out of 116 (grade 3), p=0.005 and p<0.001, respectively)

and an increased test specificity (reader 1: 91 (72.8%) out of 125 *versus* 101 (80.8%) out of 125 (grade 2) and 109 (87.2%) out of 125 (grade 3), p=0.13 and p=0.004, respectively). By contrast, in analysis 2 (fig. 3b), using the manufacturer-suggested grade 1 cut-off point, the sensitivity, specificity and LR+ for reader 1 was 56.7%, 92.2% and 7.2, respectively, and for reader 2 was 59.9%, 91.3% and 6.9, respectively. Using a grade 2 cut-off point to optimise rule-in utility, improved test specificity for both readers (reader 1: 106 (92.2%) out of 115 *versus* 113 (98.3%) out of 115, p=0.03; reader 2: 105 (91.3%) out of 115 *versus* 113 (98.3%) out of 115, p=0.02) and consequently increased LR+.

#### Performance outcomes of the LAM strip-test with and without smear microscopy

Tables 2–4 shows diagnostic accuracy measures for LAM ELISA and strip-test in all HIV-infected patients with results presented for both analysis 1 and 2, and with group 1 and 2 patients separately, and results stratified by TB diagnostic subtype or CD4 cell count. Additional tables in the online supplementary results section present data from analysis 1 and 2 separately (including LR+ values).

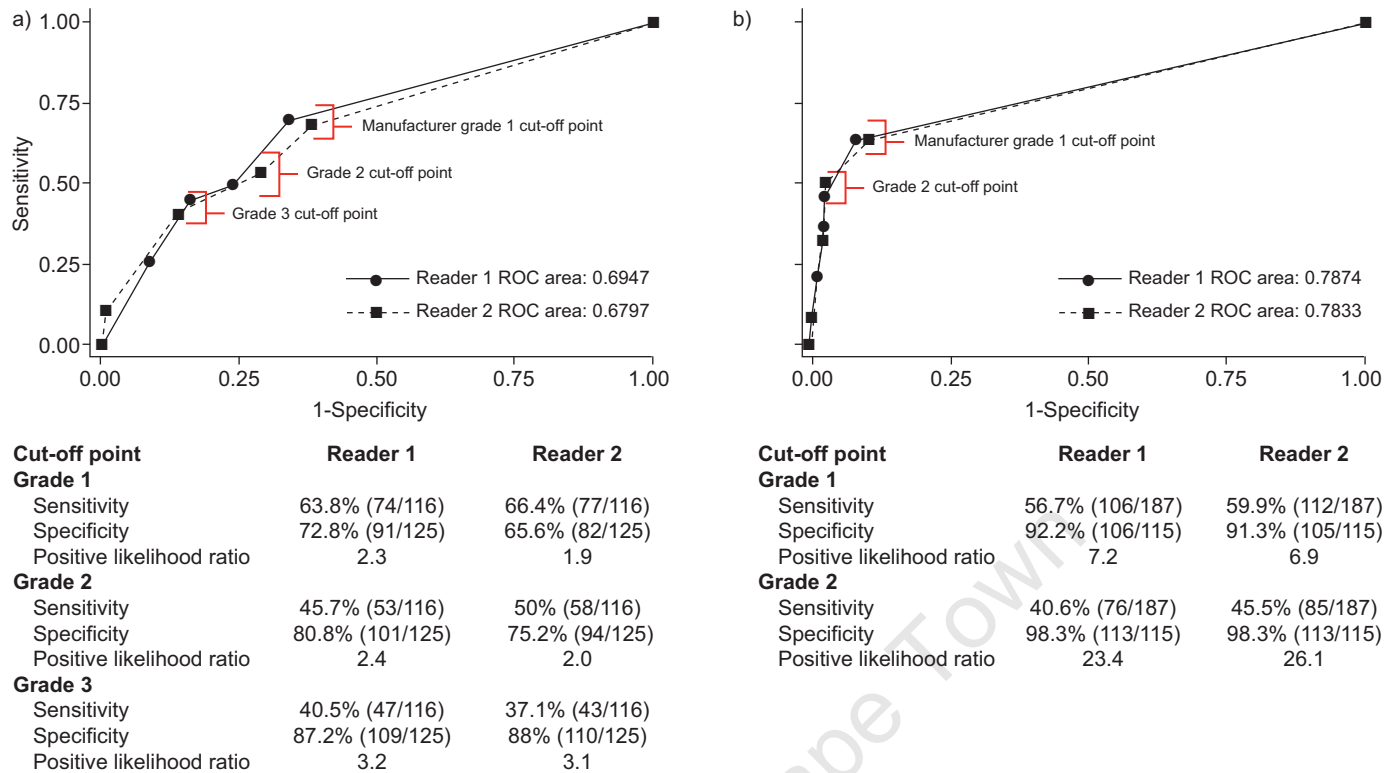
#### Analysis 1

Rapid smear microscopy had similar sensitivity to the LAM strip-test (grade 2 cut-off point) (p>0.05). LAM strip-testing (grade 2 cut-off point) was less sensitive than smear microscopy in PTB and in the combined PTB/EPTB group, but more sensitive in EPTB alone, although the numbers did not reach significance (EPTB rapid smear: 19% (8–37%) *versus* LAM strip: 41% (25–59%), p=0.07). However, if smear microscopy in the EPTB group was restricted to sputum smear microscopy only, then LAM strip-testing offered significantly improved sensitivity compared with smear microscopy: 41% (25–59%) (11 out of 27) *versus* 0% (0 out of 27), p<0.001. The sensitivity of the

**TABLE 4** Performance outcomes of smear microscopy, lipoarabinomannan (LAM) ELISA and LAM strip-test (using grade 1 and 2 cut-off points) in HIV-infected patients stratified by CD4 count (>200 and ≤200 cells per mL)

Diagnostic test(s)	HIV-infected patients with CD4 count >200 cells per mL				HIV-infected patients with CD4 count ≤200 cells per mL				
	Group 1 patients		Group 2 patients		Group 1 patients		Group 2 patients		
	Microbiological reference ( <i>M. tuberculosis</i> culture positive versus negative)	Composite reference standard (definite- and probable-TB versus non-TB)	Sensitivity	Specificity	Microbiological reference ( <i>M. tuberculosis</i> culture positive versus negative)	Composite reference standard (definite- and probable-TB versus non-TB)	Sensitivity	Specificity	
<b>Subjects n</b>	55	55	84	84	173	147	84	84	
<b>Rapid smear microscopy</b>	58 (39-75) 15/26	90 (79-100) 26/29	38 (25-52) 17/45	95 (55-100) 9/10	56 <sup>++</sup> (45-66) 45/81	91 <sup>##, \$\$</sup> (86-97) 84/92	44 <sup>##</sup> (36-53) 58/131	100 (79-100) 16/16	No results
<b>LAM ELISA</b>	27 <sup>***</sup> (14-46) 7/26	89 <sup>**</sup> (77-100) 24/27	28 <sup>***</sup> (17-43) 12/43	100 (99-100) 10/10	70 <sup>**</sup> (59-79) 56/80	76 <sup>\$\$, **</sup> (67-85) 69/91	59 <sup>***</sup> (50-67) 76/129	94 (70-100) 15/16	97 <sup>**</sup> (90-100) 28/29
<b>LAM strip-test (grade 1 cut-off point)</b>	54 (36-71) 14/26	76 (60-91) 22/29	49 (35-63) 22/45	100 (69-100) 10/10	72 (61-80) 58/81	61 <sup>#</sup> (51-71) 56/92	64 (56-72) 84/131	94 (70-100) 15/16	87 <sup>#</sup> (75-99) 26/30
<b>LAM strip-test (grade 2 cut-off point)</b>	27 <sup>†</sup> (14-46) 7/26	83 <sup>§</sup> (69-97) 24/29	29 <sup>+</sup> (18-43) 13/45	100 (69-100) 10/10	58 <sup>†, ¶</sup> (47-68) 47/81	72 <sup>†</sup> (63-81) 66/92	52 <sup>†, ++</sup> (43-60) 68/131	94 (70-100) 15/16	97 <sup>†</sup> (90-100) 29/30
<b>LAM strip-test (grade 2 cut-off point) in smear-negative/sputum scarce <i>M. tuberculosis</i> culture-positive patients</b>	0 (0-26) 0/11	N/A	14 (6-32)	N/A	51 (36-66) 21/41	N/A	43 (32-54) 31/73	N/A	N/A
<b>Combined sputum smear microscopy and LAM strip-test (grade 2 cut-off point)</b>	58 (39-75) 15/26	79 (65-94) 23/29	47 <sup>**</sup> (33-61) 21/45	95 (55-100) 9/10	75 <sup>†, ++</sup> (65-83) 61/81	70 <sup>***, ##</sup> (60-79) 64/92	68 <sup>***, ##, ++</sup> (60-75) 89/131	100 <sup>**</sup> (79-100) 16/16	No results

Data are presented as sensitivity/specificity (95% CI) n/n<sub>total</sub>, unless otherwise stated. N/A: not applicable. Using the microbiological reference, 14 out of the 242 HIV-infected group 1 patients had no CD4 cell count data, leaving 228 patients for the above analysis. 107 *Mycobacterium tuberculosis* (*M. tuberculosis*) culture-positive group 1 patients were used for the sensitivity calculations and 121 *M. tuberculosis* culture-negative group 1 for specificity calculations. Using the composite reference, 12 out of the 218 HIV-infected group 1 patients had no CD4 cell count data, leaving 202 patients for the above analysis. 176 definite- and probable-tuberculosis (TB) were used for the sensitivity calculations and 26 non-TB patients for specificity calculations. Additionally, 84 non-TB group 2 patients were used for a second specificity calculation. Rapid (within 24 h of hospitalisation) smear microscopy includes both sputum and non-sputum samples (e.g. fine-needle aspirate and cerebrospinal fluid samples). Two CD4 >200 cells per mL and CD4 ≤200 cells per mL group 1 patients (analysis 1) and three CD4 >200 and CD4 ≤200 (analysis 2) had a missing LAM ELISA result while one CD4 >200 (analysis 2) had a missing LAM strip-test result. For a comparison of differences in diagnostic accuracy measures between analyses and patient groups for a specific test (LAM ELISA or LAM strip-test) p-values were p<0.05; specific p-values: \*\*; p = 0.009; †; p = 0.006; ‡; p = 0.008; §; p = 0.002; ¶; p = 0.004; nonsignificant p-values are not shown. For a comparison of the specific measures of diagnostic accuracy (sensitivity or specificity) between different tests (LAM ELISA versus LAM strip-test) or combinations thereof, or between different cut-off points of the urine LAM strip-test (grade 1 versus grade 2 cut-off point) p-values were p<0.05; specific p-values: ##; p = 0.002; ††; p = 0.008; †††; p = 0.005; nonsignificant p-values not shown.



**FIGURE 3.** Receiver operator characteristic (ROC) curves for urine lipoarabinomannan strip-test graded by two independent readers for HIV-infected patients. Sensitivity, specificity and positive likelihood ratios are shown for different visual band intensity-based cut-off points (grade 1 (manufacturer recommended) and grades 2 and 3 (optimised for “rule-in”) are indicated). a) Analysis 1 (*Mycobacterium tuberculosis* culture reference standard). b) Analysis 2 (composite reference standard).

LAM strip-test (grade 2 cut-off point) in smear negative/sputum scarce HIV-infected TB patients was 39% (28–52%), improving to 51% (36–66%) in HIV-infected patients with a CD4 cell count  $\leq 200$  cells per mL, and 58% (41–74%) when the CD4 cell count was  $\leq 100$  cells per mL ( $p=0.2$  and  $p=0.09$ , respectively; see online supplementary tables for CD4  $\leq 100$  cells per mL results). Overall, the sensitivity (95% CI) of the LAM strip-test (grade 2 cut-off point) was higher in CD4  $\leq 200$  cells per mL versus CD4  $> 200$  cells per mL (58% (47–68%) versus 27% (14–46%),  $p=0.006$ ). No significant differences were noted for the above tests amongst patients with CD4  $\leq 200$  cells per mL,  $\leq 100$  cells per mL and  $\leq 50$  cells per mL (data not shown). The combined performance of rapid smear microscopy and the urine LAM strip-test (grade 2 cut-off point) in HIV-infected patients with a CD4 cell count  $\leq 200$  cells per mL was significantly better than either sputum smear microscopy or the LAM strip-test alone (75% (65–83%) versus 56% (45–66%) versus 58% (47–68%),  $p=0.008$  and  $p=0.02$ ). TB LAM ELISA results were not significantly different to LAM strip-test results at both grade 1 and 2 cut-off points. The specificity of the TB LAM ELISA and LAM strip-test (both cut-off points) in group 1 patients was 80% (72–86%) and 75% (67–82%), respectively, while in group 2 patients, specificity was 99% (94–100%) for both the LAM ELISA and strip-tests. In both CD4  $> 200$  cells per mL and  $\leq 200$  cells per mL LAM strip-test (grade 2 cut-off point), specificity was higher in group 2 than in group 1 patients (CD4  $> 200$  cells per mL: 100% (93–100%) versus 83% (69–97%),  $p=0.002$ ; CD4  $\leq 200$  cells per mL: 97% (90–100%) versus 72% (63–81%),  $p=0.004$ ). In group 2 patients, LAM strip-test specificity was

higher using the grade 2 versus grade 1 cut-off point (100% (93–100%) versus 91% (83–99%),  $p=0.02$ ).

#### Analysis 2

As in analysis 1, rapid smear microscopy had similar sensitivity to the LAM strip-test (grade 2 cut-off point) ( $p>0.05$ ). Overall, the sensitivity of the TB LAM ELISA and LAM strip-test (both cut-off points) showed a small, but nonsignificant reduction compared with analysis 1, but improvements between patient subgroups were similar between analyses. The sensitivity of the LAM strip-test (grade 2 cut-off point) in smear negative/sputum scarce HIV-infected TB patients was 38% (29–48%), improving to 43% (32–54%) in HIV-infected patients with a CD4 cell count  $\leq 200$  cells per mL, and 49% (35–63%) when the CD4 cell count was  $\leq 100$  cells per mL ( $p=0.5$  and  $p=0.2$ , respectively; see online supplementary tables for CD4  $\leq 100$  cells per mL results). The sensitivity (95% CI) of the LAM strip-test (grade 2 cut-off point) was higher in CD4  $\leq 200$  cells per mL versus CD4  $> 200$  cells per mL (52% (43–60%) versus 29% (18–43%),  $p=0.008$ ). The combined performance of rapid smear microscopy and the urine LAM strip-test (grade 2 cut-off point) in HIV-infected patients with a CD4 cell count  $\leq 200$  cells per mL, was significantly better than either sputum smear microscopy or the LAM strip-test alone (68% (60–75%) versus 44% (36–53%) versus 52% (43–60%),  $p<0.001$  and  $p=0.008$ ). Notably, the overall specificity of both the TB LAM ELISA and LAM strip-test (both cut-off points) was higher in analysis 2 than in analysis 1 (LAM strip-test (grade 2 cut-off point) specificity: analysis 2: group 1 96% (89–100%) and group 2 99% (94–100%) versus

analysis 1: 75% (67–82%),  $p=0.01$  and  $p<0.001$ , respectively). A similar difference in test specificities between analysis 1 and 2 was found in patients with  $CD4 >200$  cells per mL and  $\leq 200$  cells per mL (LAM strip-test (grade 2 cut-off point) specificity in  $CD4 \leq 200$  cells per mL: analysis 2: group 1 94% (70–100%) and group 2 97% (90–100%) versus analysis 1: 72% (63–81%),  $p=0.06$  and  $p=0.004$ , respectively).

#### Factors associated with a positive LAM strip-test result

A positive LAM strip-test was associated with a  $CD4$  cell count  $\leq 200$  cells per mL ( $p<0.001$  for both reader 1 and 2), but not with any other clinical parameter including urinalysis results, renal function results (urea and creatinine levels) and/or calculated glomerular filtration rate (using the modified Cochrane–Gault method).

#### LAM strip-test inter- and intraobserver variability

At the grade 1 cut-off point, the inter- and intraobserver agreement between readers was fair to good ( $\kappa=0.45$ – $0.59$  for interobserver agreement and  $0.56$ – $0.73$  for intraobserver agreement). By contrast, at the grade 2 cut-off point, the inter- and intraobserver agreement improved and was moderate to very good ( $\kappa=0.78$ – $0.92$  for interobserver agreement and between  $\kappa=0.92$ – $0.96$  for intraobserver agreement). Interobserver agreement improved significantly by using a grade-2 versus a grade-1 cut-off point ( $p=0.002$ ). Detailed tables of the agreement between all the readers for different LAM strip-test grades is provided in the online supplement results section.

#### LAM strip-test indeterminate tests

423 LAM strip-tests were performed and read by two independent readers. The initial indeterminate rate (one test) for the LAM strip-test (defined as a broken band in the patient window) was 29–60 (7–14%) out of 423 depending on the reader. In  $>90\%$  of these cases, a repeat LAM strip-test produced a valid result. Thus, only 1–2% of tests remained indeterminate after repeat testing. There was no significant difference if performance outcomes were calculated excluding indeterminate test results or considering them as “negative” test results. Given the proposed use of the LAM strip-test as a rapid rule-in test, the analyses presented above consider indeterminate test results as test “negative”. Three failed tests occurred (no band visible in the control window) and they were successfully repeated.

## DISCUSSION

In Africa, HIV-TB co-infected patients form the bulk of medical admissions, have prolonged hospital stay due to diagnostic delay, have a high mortality if treatment is not rapidly initiated and overburden public health facilities. To address these issues, we evaluated the accuracy and clinical utility of a urine POC LAM strip-test to rule-in TB in HIV-infected hospitalised adults in a resource-poor high HIV-prevalence setting. The key findings of our study were that in hospitalised HIV co-infected patients: 1) the LAM strip-test had inadequate specificity at the manufacturer’s recommended cut-off point; 2) using the grade 2 cut-off point, LAM offers better reliability (minimal interobserver variability) and utility as a rule-in test; 3) LAM strip-testing offered incremental value over smear microscopy in EPTB and thus, if used in combination with smear microscopy for HIV-infected hospitalised patients, nearly three-quarters of definite-TB cases were identifiable; and 4) the LAM strip-test was positive

in approximately one-half of all sputum-scarce or smear-negative TB HIV-infected patients with a  $CD4 \leq 200$  cells per mL who would have otherwise required further investigation.

Importantly, the LAM strip-test was particularly useful in patients with EPTB and those with a  $CD4$  count  $\leq 100$  cells per mL, potentially offering a rapid rule-in diagnosis in these diagnostically challenging patient subgroups. Thus, a convenience urine sample could be obtained in patients with PTB and/or EPTB irrespective of whether the patient could produce sputum. It is in this context, that the LAM assay may also offer advantage over other novel diagnostic formats, such as the Xpert MTB/RIF, as it obviates the need for a biological sample from the site of disease. Thus, although the assay is useful in only a specific patient subgroup, we believe that these results are encouraging given that the urine LAM strip-test is a simple-to-perform bedside test that is affordable (less than US\$3.50 per test), rapid (providing results within 35 min) and uses an easily obtainable urine sample. These factors make it ideally suited for use in hospitals in high-burden TB–HIV settings where resources and laboratory infrastructure are limited.

What is an optimal and pragmatic cut-off point to select when using the LAM strip-test? The manufacturer recommends the grade 1 cut-off point (fig. 1), and this cut-off point proved reliable and specific in a recently published study in an outpatient antiretroviral-screening clinic [19]. However, in our study, using the grade 1 cut-off point the LAM strip-test had suboptimal specificity (in both analyses) and modest interobserver agreement, thereby limiting rule-in test utility. Moreover, it was noted that, when the faintest bands were being interpreted in differing ambient light conditions, readings were variable. By contrast, using the grade 2 cut-off point was considerably more reliable and improved test specificity, and is hence the one that we recommend. The reasons for differing cut-off point-specific performance differences between the data of LAWN *et al.* [19] and our data remains unclear, but may be related to within batch variability, impact of storage duration and conditions, and reader blinding and experience. Further prospective standardised studies are now urgently required to clarify these findings and the optimal rule-in cut-off point in different settings.

Given the known limitations of TB culture as a reference standard, especially amongst HIV-infected hospitalised TB patients (very low sputum bacillary load and sputum scarce), we performed two analyses of diagnostic accuracy using two reference standards. Furthermore, a control group of HIV-infected non-TB patients was enrolled to obtain an additional estimate of test specificity. Specificity measures showed significant variation between different patient groups and analyses, with analysis 1 likely to provide an underestimate and analysis 2 an overestimate of test specificity. We feel that, given this particular patient population, and the large number of probable (likely) TB cases in the cohort (including smear-positive culture negative patients), analysis 2 test specificity is a more representative and accurate. Larger studies using a standardised TB diagnostic evaluation are now urgently required to clarify our findings.

Our study has a number of limitations. For practical reasons, our study did not use standardised TB evaluation, and diagnostic work-up was at the discretion of the attending clinician. This may have introduced work-up bias. Our study used frozen samples to

evaluate the LAM strip-test and this may have impacted test performance. However, sample freeze-thaw showed no significant effect in our preliminary experiments or in recent meta-analysis of studies using the TB LAM ELISA [13]. The LAM strip-test does not inform on drug susceptibility and thus additional samples may be required, if appropriate, to make a diagnosis of multi-drug resistant (MDR)-TB. This consideration, hitherto of limited general applicability, may become more relevant in the future when National TB Programmes in Africa develop the capacity to diagnose and treat MDR-TB. Finally, LAM test results were not used for treatment decisions or to guide further diagnostic work-up and, consequently, the clinical impact on patient mortality, time to diagnosis, length of hospital stay and cost-effectiveness could not be evaluated.

The novel LAM strip-test, although restricted to HIV-infected persons with advanced immunosuppression, targets a vulnerable population with high mortality and healthcare-related costs. Given its low cost and ease of use, it has the potential to offer significant clinical utility when combined with smear microscopy in high-burden district hospitals with poor or no laboratory infrastructure. If used to guide the rapid initiation of anti-TB treatment, the impact of which is supported by a recent study in HIV-infected hospitalised patients [4], LAM strip-testing has the potential to decrease TB-related mortality, length of hospital stay and the overall cost of the TB diagnostic work-up in resource-poor healthcare systems. Suboptimal specificity may limit this potential clinical utility. Larger prospective studies of diagnostic accuracy, using a standardised TB diagnostic evaluation, are now urgently required to clarify test specificity in hospitalised patients with TB-HIV co-infection and confirm our recommendation to use the grade-2 cut-off point for optimising rule-in value and inter-reader reliability.

#### SUPPORT STATEMENT

Although Alere supplied the ELISA and LAM strip-tests free of charge they had no role in the design and conduct of the study, analysis of the data or writing of the manuscript. J.G. Peter and R. van Zyl-Smit are supported by the Fogarty International Clinical Research Scholars/Fellows Support Centre National Institutes of Health grant R24TW007988, SATBAT and the EDCTP. G. Theron is supported by the EDCTP and NRF. K. Dhedda is supported by the EDCTP (TB NEAT and TESA) and the SA DST and NRF (SARChI).

#### STATEMENT OF INTEREST

A statement of interest for the study can be found at [www.erj.ersjournals.com/site/misc/statements.xhtml](http://www.erj.ersjournals.com/site/misc/statements.xhtml)

#### ACKNOWLEDGEMENTS

We thank the research nursing staff, including N. Kelly and F. Embrahim (University of Cape Town, Cape Town, South Africa), who were involved in the recruitment of patients and collection of urine samples. The support of the clinical and nursing staff of the general medical wards of Groote Schuur, Victoria, G.F. Jooste and New Somerset Hospitals (all Cape Town, South Africa) was greatly appreciated.

#### REFERENCES

- 1 WHO. Global tuberculosis control 2010. WHO/HTM/TB/2010.7. Geneva, World Health Organisation, 2010.
- 2 Getahun H, Gunneberg C, Granich R, *et al.* HIV infection-associated tuberculosis: the epidemiology and the response. *Clin Infect Dis* 2010; 50: Suppl. 3, S201–S207.
- 3 Harrison D. An overview of health and health care in South Africa 1994–2010: priorities, progress and prospects for new gains; 2009. Online report.
- 4 Holtz TH, Kabera G, Mthiyane T, *et al.* Use of a WHO-recommended algorithm to reduce mortality in seriously ill patients with HIV infection and smear-negative pulmonary tuberculosis in south africa: an observational cohort study. *Lancet Infect Dis* 2011; 11: 533–540.
- 5 Cohen T, Murray M, Wallengren K, *et al.* The prevalence and drug sensitivity of tuberculosis among patients dying in hospital in kwazulu-natal, South Africa: a postmortem study. *PLoS Med* 2010; 7: e1000296.
- 6 Cain KP, McCarthy KD, Heilig CM, *et al.* An algorithm for tuberculosis screening and diagnosis in people with HIV. *N Engl J Med* 2010; 362: 707–716.
- 7 Monkongdee P, McCarthy KD, Cain KP, *et al.* Yield of acid-fast smear and mycobacterial culture for tuberculosis diagnosis in people with human immunodeficiency virus. *Am J Respir Crit Care Med* 2009; 180: 903–908.
- 8 Van Rie A, Page-Shipp L, Scott L, *et al.* Xpert(®) MTB/RIF for point-of-care diagnosis of TB in high-HIV burden, resource-limited countries: hype or hope? *Expert Rev Mol Diagn* 2010; 10: 937–946.
- 9 Mishra AK, Driessen NN, Appelmeik BJ, *et al.* Lipoarabinomannan and related glycoconjugates: structure, biogenesis and role in *Mycobacterium tuberculosis* physiology and host-pathogen interaction. *FEMS Microbiol Rev* 2011; 35: 1126–1157.
- 10 Hamasur B, Bruchfeld J, Haile M, *et al.* Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine. *J Microbiol Methods* 2001; 45: 41–52.
- 11 Tessema TA, Bjune G, Hamasur B, *et al.* Circulating antibodies to lipoarabinomannan in relation to sputum microscopy, clinical features and urinary anti-lipoarabinomannan detection in pulmonary tuberculosis. *Scand J Infect Dis* 2002; 34: 97–103.
- 12 Peter J, Green C, Hoelscher M, *et al.* Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. *Curr Opin Pulm Med* 2010; 16: 262–270.
- 13 Minion J, Leung E, Talbot E, *et al.* Diagnosing tuberculosis with urine lipoarabinomannan: systematic review and meta-analysis. *Eur Respir J* 2011; 38: 1398–1405.
- 14 Gounder CR, Kufa T, Wada NI, *et al.* Diagnostic accuracy of a urine lipoarabinomannan enzyme-linked immunosorbent assay for screening ambulatory HIV-infected persons for tuberculosis. *J Acquir Immune Defic Syndr* 2011; 58: 219–223.
- 15 Flores L, Steingart K, Dendukuri N, *et al.* Systematic review and meta-analysis of antigen detection tests for the diagnosis of tuberculosis. *Clin Vaccine Immunol* 2011; 18: 1616–1627.
- 16 Shah M, Variava E, Holmes CB, *et al.* Diagnostic accuracy of a urine lipoarabinomannan test for tuberculosis in hospitalized patients in a high HIV prevalence setting. *J Acquir Immune Defic Syndr* 2009; 52: 145–151.
- 17 Lawn SD, Edwards DJ, Kranzer K, *et al.* Urine lipoarabinomannan assay for tuberculosis screening before antiretroviral therapy: diagnostic yield and association with immune reconstitution disease. *AIDS* 2009; 23: 1875–1880.
- 18 Alere. Alere Determine TB LAM Ag rapid test product listing. 2011. [www.alerehiv.com/hiv-comorbidities/features](http://www.alerehiv.com/hiv-comorbidities/features)
- 19 Lawn SD, Kerkhoff AD, Vogt M, *et al.* Diagnostic accuracy of a low-cost, urine antigen, point-of-care screening assay for HIV-associated pulmonary tuberculosis before antiretroviral therapy: a descriptive study. *Lancet Infect Dis* 2012; 12: 201–209.
- 20 Dheda K, Davids V, Lenders L, *et al.* Clinical utility of a commercial LAM-ELISA assay for TB diagnosis in HIV-infected patients using urine and sputum samples. *PLoS One* 2010; 5: e9848.
- 21 Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159–174.
- 22 Bossuyt PM, Reitsma JB, Bruns DE, *et al.* The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. The standards for reporting of diagnostic accuracy group. *Croat Med J* 2003; 44: 639–650.

## Urine antigen test for diagnosis of HIV-associated tuberculosis

The need for rapid, affordable, bedside tests for tuberculosis is a global priority.<sup>1</sup> Stephen Lawn and colleagues presented a study<sup>2</sup> assessing the urine Determine TB-LAM test for tuberculosis screening in HIV-infected patients before the initiation of antiretrovirals. Their findings of moderate test sensitivity but excellent specificity, potential low cost, and usefulness when combined with sputum smear microscopy in HIV-infected patients with advanced immunosuppression were interesting and encouraging. The test is thus potentially useful in these diagnostically challenging patients, who have high tuberculosis-related morbidity.<sup>2-4</sup> We assessed the

same test in hospital inpatients with advanced immunosuppression co-infected with HIV and tuberculosis and, by contrast with the findings of Lawn and colleagues, recorded notable differences in interobserver agreement and test specificity.<sup>5</sup> In view of the potential point-of-care use of this test, we believe that our findings should be emphasised.

In both studies, the test was read at the manufacturer's suggested grade 1 cut-point under similar conditions (figure); however, Lawn and colleagues noted substantially better interobserver proportional agreement than we did (99.6% vs 89.1%,  $p < 0.0001$ ).<sup>2,5</sup> Inter-reader disagreement in our study was due to the variability and misinterpretation of very faint (grade 1) bands. Agreement improved when a grade 2 cut-point was used (95.1% vs 89.1%,  $p = 0.002$ ).<sup>5</sup> Additionally, use of a

grade 2 cut-point improved test specificity in a healthy control group (99% vs 90%,  $p = 0.009$ ).<sup>2</sup>

The reason for these population-specific differences is unclear. The 24 positive results with Determine TB-LAM reported by Lawn and colleagues<sup>2</sup> had high corresponding lipoarabinomannan ELISA optical densities. By contrast, in our cohort of hospital inpatients with higher tuberculosis prevalence and lower median CD4 cell counts than Lawn's cohort, low ELISA optical densities and corresponding faint bands were frequent. Within-batch variability and differences in reader experience might have been contributing factors.

The performance and agreement of the strip test thus seems to be setting-specific. Caution should be applied when generalising the findings of Lawn and colleagues to other settings. Further studies are urgently needed to delineate context and setting-specific performance and optimise cut-point selection for this potentially promising point-of-care test.

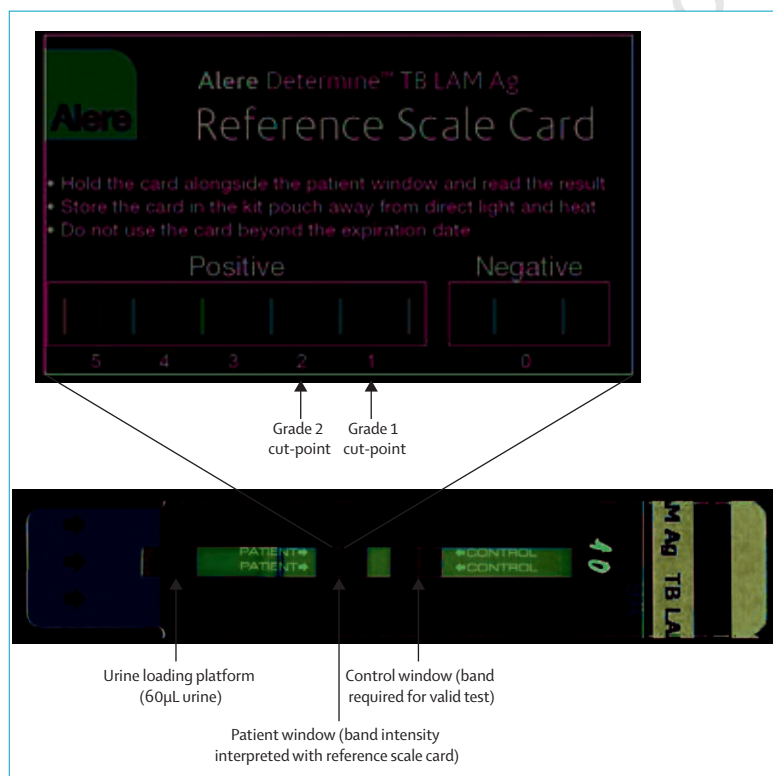
Alere donated LAM ELISA-kits and strip tests. We declare that we have no conflicts of interest.

Jonathan G Peter, Grant Theron,  
\*Keertan Dheda

keertan.dheda@uct.ac.za

Lung Infection and Immunity Unit, Department of Medicine, University of Cape Town, Cape Town 7925, South Africa

- 1 Keeler E, Perkins MD, Small P, et al. Reducing the global burden of tuberculosis: the contribution of improved diagnostics. *Nature* 2006; **444** (suppl 1): 49–57.
- 2 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.
- 3 Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, Dheda K. Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. *Curr Opin Pulm Med* 2010; **16**: 262–70.
- 4 Dheda K, Davids V, Lenders L, et al. Clinical utility of a commercial LAM-ELISA assay for TB diagnosis in HIV-infected patients using urine and sputum samples. *PLoS One* 2010; **5**: e9848.
- 5 Peter JG, Theron G, van Zyl Smit R, et al. Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients. *Eur Respir J* 2012; published online Feb 23. DOI:10.1183/09031936.00201711.



**Figure:** Determine TB-LAM test and reference scale card

The reference scale card shows six cut-points (visual grades 0–5) categorised by different band intensities in the patient window. The manufacturers suggested grade-1 cut-point and also the grade-2 cut-point are highlighted.

## Chapter 7.

The diagnostic accuracy of urine-based Xpert MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce\*

Peter, J.G., Theron, G., Muchinga, T.E., Govender, U. and Dheda, K.

Published in *PLoS ONE* 7(7): e39966.

\*asterisked and light gray sections of this chapter form part of a recent MSc submission to the University of Cape Town for MSc candidate T.E. Muchinga. I was the co-supervisor of this candidates MSc work and submission.

### PhD Context (Chapters 7 & 8)

Urinary LAM, either used as a strip test or in the original ELISA format, offers modest diagnostic sensitivity for SN- and SSTB in HIV-infected hospitalised patients, and the use of other alternative urine-based diagnostic tools may improve sensitivity. Both another research group and our group have hypothesised that urine LAM positivity may be associated with the presence of a “leaky” glomerular filtration membrane and mycobacteriuria (as an extension of the known association of increased urine LAM with high total mycobacterial burden). Consequently, we decided to test the urine samples from our prospective cohort of patients (described in detail in Chapter 6), which had already been tested with the urinary LAM ELISA and strip tests, with the MTB/RIF assay. Conducting this line of investigation allowed us to: i) explore our hypothesis of the determinates of urine LAM positivity; ii) evaluate and optimise the performance of MTB/RIF on hitherto unevaluated biological sample, namely urine; iii) determine

if the performance of urine-based MTB/RIF could offer any incremental diagnostic utility, alone or in combination with urine LAM, in SN- or SSTB; and iv) assess the relationship between urine LAM concentration and urine-based MTB/RIF-generated CT-value (described in Chapter 3 to correlate excellently with smear-grade and time-to-culture positivity).

University of Cape Town

# The Diagnostic Accuracy of Urine-Based Xpert MTB/RIF in HIV-Infected Hospitalized Patients Who Are Smear-Negative or Sputum Scarce

Jonathan G. Peter<sup>1</sup>, Grant Theron<sup>1</sup>, Tapuwa E. Muchinga<sup>1</sup>, Ureshnie Govender<sup>1</sup>, Keertan Dheda<sup>1,2,3\*</sup>

**1** Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, University of Cape Town, Cape Town, South Africa, **2** Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa, **3** Department of Infection, University College London Medical School, London, United Kingdom

## Abstract

**Background:** Hospitals in sub-Saharan Africa are inundated with HIV-infected patients and tuberculosis (TB) is the commonest opportunistic infection in this sub-group. Up to one third of TB-HIV co-infected patients fail to produce a sputum sample (sputum scarce) and diagnosis is thus often delayed or missed. We investigated the sensitivity of urine-based methods (Xpert MTB/RIF, LAM strip test and LAM ELISA) in such patients.

**Methodology/Principal Findings:** 281 HIV-infected hospitalised patients with clinically suspected TB provided a spot urine sample. The reference standard was culture positivity for *Mycobacterium tuberculosis* on  $\geq 1$  sputum or extra-pulmonary sample. MTB/RIF was performed using 1 ml of both unprocessed and, when possible, concentrated urine. Each unconcentrated urine sample was also tested using the Clearview LAM ELISA and Alere LAM strip test. 42% (116/242) of patients had culture-proven TB. 18% (20/54) were sputum scarce. In sputum-scarce patients, the sensitivity of urine MTB/RIF and LAM ELISA was 40% (95%CI: 22–61) and 60% (95%CI: 39–78), respectively. Urine MTB/RIF specificity was 98% (95%CI: 95–100). Combined sensitivity of urine LAM ELISA and MTB/RIF was better than MTB/RIF alone [MTB/RIF and LAM: 70% (95%CI: 48–85) vs. MTB/RIF: 40% (95%CI: 22–61),  $p=0.03$ ]. Significant predictors of urine MTB/RIF positivity were CD4<50 cells/ml ( $p=0.001$ ), elevated protein-to-creatinine ratio ( $p<0.001$ ) and LAM ELISA positivity ( $p<0.001$ ). Urine centrifugation and pelleting significantly increased the sensitivity of MTB/RIF over unprocessed urine in paired samples [42% (95%CI: 26–58) vs. 8% (95%CI: 0–16),  $p<0.001$ ]. Urine MTB/RIF-generated  $C_T$  values correlated poorly with markers of bacillary burden (smear grade and time-to-positivity).

**Conclusions/Significance:** This preliminary study indicates that urine-based MTB/RIF, alone or in combination with LAM antigen detection, may potentially aid the diagnosis of TB in HIV-infected patients with advanced immunosuppression when sputum-based diagnosis is not possible. Concentration of urine prior to MTB/RIF-testing significantly improves sensitivity.

**Citation:** Peter JG, Theron G, Muchinga TE, Govender U, Dheda K (2012) The Diagnostic Accuracy of Urine-Based Xpert MTB/RIF in HIV-Infected Hospitalized Patients Who Are Smear-Negative or Sputum Scarce. PLoS ONE 7(7): e39966. doi:10.1371/journal.pone.0039966

**Editor:** Delia Goletti, National Institute for Infectious Diseases (L. Spallanzani), Italy

**Received:** March 8, 2012; **Accepted:** June 5, 2012; **Published:** July 9, 2012

**Copyright:** © 2012 Peter et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** This project was funded by a grant from the European and Developing Countries Clinical Trials Partnership (EDCTP). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** Keertan Dheda is a Plos ONE editorial board member. This does not alter our adherence to all the PLoS ONE policies on sharing data and materials.

\* E-mail: Keertan.Dheda@uct.ac.za

## Introduction

In Africa, up to 65% of active tuberculosis (TB) cases are co-infected with HIV [1]. TB-related mortality is highest in this patient sub-group, and district-level hospitals are inundated with patients with advanced immunosuppression. With advancing HIV-related immunosuppression, the frequency of extra-pulmonary (EPTB) and disseminated forms of TB disease increase [2,3], sputum smear microscopy performance is reduced, and up to a third of patients are unable to produce sputum for diagnostic testing [4]. Diagnosis is therefore challenging and often delayed, and post-mortem studies reveal a large burden of undiagnosed TB in HIV-infected hospitalised patients [5,6,7]. Recent studies have indicated that the rapid initiation of anti-TB treatment may

reduce mortality [8]. There is a clear need for new, accurate, and rapid TB diagnostics that have utility in patients who cannot produce sputum.

The Clearview TB LAM ELISA (Alere Medical innovations, USA) detects LAM antigen in the urine and has recently evolved into a new point-of-care lateral flow test (Alere Determine-TB LAM Ag strip test) [9]. We recently found that this assay offered the greatest benefit in hospitalised HIV co-infected patients with advanced immunosuppression [4]. By contrast, the MTB/RIF assay is a novel, automated molecular TB diagnostic able to detect both the presence of *Mycobacterium tuberculosis* complex DNA and rifampicin drug-resistance in less than two hours. This test has been endorsed by the World Health Organization and is being rolled out in South Africa as a frontline test for individuals with

suspected TB [10,11]. Given the high accuracy of this test in sputum samples (sensitivity and specificity of 90% and 99%) [12], it represents a considerable advance over smear microscopy for the diagnosis of pulmonary-TB. However, acquiring a diagnostic sample remains a major hurdle in HIV-infected sputum scarce patients suspected of having active TB. Sputum induction, using ultrasonic nebulisation, may facilitate obtaining sputum, but this is often unavailable in hospitals in resource-poor settings and infection control is a concern. Tissue biopsies and aspirated samples may be obtained from extra-pulmonary disease foci (e.g. bone marrow and liver, pleural and pericardial fluid) but specialised skill and equipment requirements limit the availability and affordability in resource-poor settings. Urine is easily obtainable from sputum scarce patients but there are few data about the performance of newer diagnostic tests using urine [13].

We hypothesised that urine MTB/RIF may offer diagnostic utility in patients where sputum-based diagnosis is not feasible. The performance of this test specifically in sputum scarce patients with HIV has not been previously evaluated using urine.

## Methods

### Study Population

The study population consisted of 335 prospectively recruited adult patients from four hospitals (three district- and one tertiary-level) between July 2009 and December 2010 in Cape Town, South Africa. Patients were referred for study inclusion by attending clinicians if the patient was suspected to have HIV-TB co-infection. Only three patients, who refused consent, were excluded from study enrolment. All other patients provided written informed consent and the study was approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee. Clinical information documented for enrolled patients included demographic information, past history of TB, co-morbidity, symptoms and vital signs, HIV status and renal function. A study outline is shown in Figure 1.

### Diagnostic Sample Collection Handling

Consultant-led groups of attending clinicians with no association to the study team decided on the timing and extent of diagnostic work-up, commencement of empiric anti-TB treatment, and final discharge from hospital. TB diagnostic work-up was not standardised, but routine local hospital practice includes the collection of two sputum samples in patients able to expectorate and, if EPTB is suspected, the collection of 1–2 non-sputum samples from clinically involved sites (e.g. fine needle aspirate of lymph node, pleural fluid aspirate/biopsy, ascitic tap, lumbar puncture, pericardial aspiration etc.). Further details of biological samples collected for TB culture are provided in table 1. The local reference laboratory processed all clinical specimens collected for TB diagnosis. Fluorescence smear microscopy was performed on NALC/NaOH processed sputum, which was also cultured using the MGIT 960 liquid culture system (BD Diagnostics, USA). The reference standard for definite-TB was liquid culture positivity for *M. tuberculosis*.

### Urine Sampling and LAM Methodology

All patients were required to give a spot urine sample (10–30 ml) collected in a sterile container as soon as possible after recruitment. A urine dipstick test (*UriCHECK* 9, RapiMed Diagnostics, South Africa) was immediately performed to assess for protein, blood and leucocytes, while the local reference laboratory performed urinary protein and creatinine measurements. Urine was frozen on the day of collection and stored at

–20°C for later batched testing. Both the LAM ELISA and LAM strip test (a single manufacturing lot#101102) were performed on thawed urine according to the manufacturers' instructions by readers blinded to patient data. Detailed methodology for both tests has been previously described [14]. Of note is the use of the grade 2 cut-point and not the manufacturer's suggested grade 1 cut-point for the LAM strip test. This cut-point optimises test specificity and rule-in value [4] in hospitalised patients with HIV.

### Urine MTB/RIF Methodology

All HIV-infected patients with culture positive TB had an MTB/RIF performed using 1 ml of unprocessed, thawed urine according to the manufacturers' suggested procedure for sputum samples [15]. In addition, a random sample of ~50% (62/126) of culture negative non-TB patients had a urine MTB/RIF performed. The MTB/RIF operator was blinded to the clinical status of these patients. Briefly, the sample reagent was mixed at a 2:1 ratio with ~1 ml of urine. Two millilitres of the reagent sample mix was transferred into an MTB/RIF assay cartridge and inserted into the GeneXpert instrument [15]. Additionally, if the MTB/RIF was negative using a 1 ml urine sample, a second pelleted urine MTB/RIF was performed, where possible, using a median (IQR) of 10 (5–10) ml urine. Urine was centrifuged at 3000 g for 15 minutes and the pellet re-suspended in 1 ml of sterile phosphate-buffered saline. In culture-negative non-TB urine samples used for MTB/RIF, pelleting of up to 10 mls was performed where possible.

### Statistical Analysis

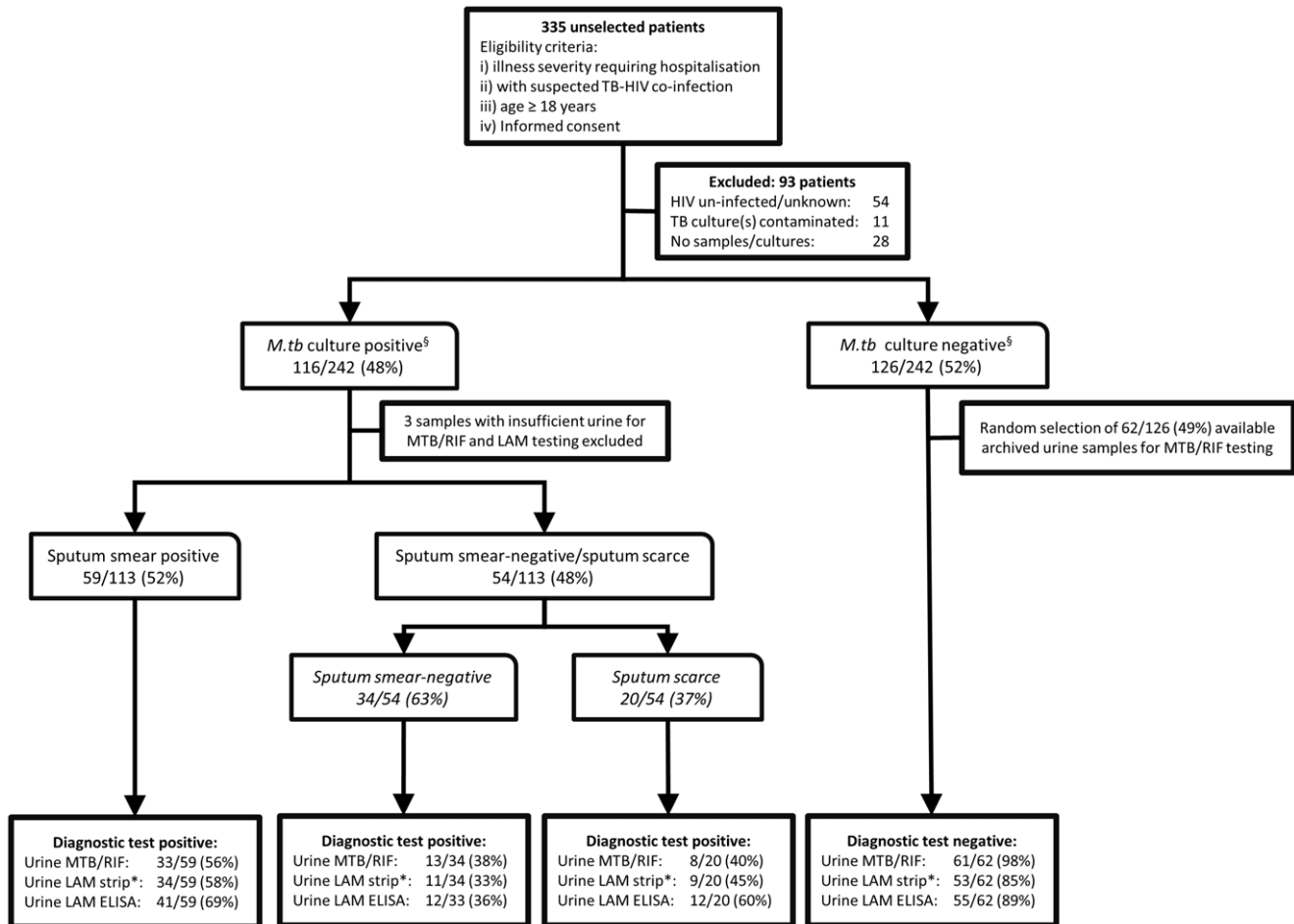
Sensitivity and specificity measures for all diagnostic tests are presented with 95% confidence intervals. Demographic, clinical and microbiological characteristics of different patient sub-groups were compared using  $\chi^2$  and Wilcoxon rank-sum test as appropriate. Diagnostic sensitivity and specificity of individual and/or combinations of tests was compared using the  $\chi^2$  and Fisher's exact tests, as appropriate. Logistic regression analysis was used to identify predictors of urine MTB/RIF positivity (in all patients and restricted to smear-negative and sputum scarce patients). Spearman correlation ( $R_s$ ) was used to evaluate relationship between MTB/RIF-generated  $C_T$  values and other markers of bacillary load. All statistical tests were 2 sided at  $\alpha=0.05$ . STATA IC, version 10 (Stata Corp, Texas, USA) was used for all statistical analyses. Study reporting and analysis were consistent with the STARD criteria [16].

## Results

### Study Population and Proportion of Sputum Scarce Patients

Figure 1 outlines the study population and test results. 93 patients were excluded from the primary analysis, leaving 242 patients with  $\geq 1$  sputum or non-sputum (or both) liquid culture result. A further three patients had insufficient urine for MTB/RIF, LAM ELISA and strip testing. Patient demographic and clinical characteristics stratified by sputum smear status and urine MTB/RIF results are shown in Table 1. Additionally, table 1 shows details of all biological samples undergoing mycobacterial liquid culture. 48% (116/242) of included patients had culture-positive TB from either a sputum ( $n=68$ ), non-sputum sample ( $n=26$ ), or both ( $n=22$ ). Only 1/113 of culture positive patients had a positive urinary *m.tb* culture.

17% (42/242) of all patients with valid culture results and 18% (20/113) of culture positive were unable to produce sputum



**Figure 1. Study flow diagram.** <sup>§</sup>Both sputum and non-sputum samples (e.g. blood, pleural or pericardial fluid) were collected by attending clinicians for liquid TB culture. *M.tb* culture positive patients had at least one sputum or non-sputum sample liquid culture positive, while *m.tb* culture negative patients had at least one (usually 2 or more) samples liquid culture negative \*Results using a grade 2 cut-point for the urine LAM strip test are shown. doi:10.1371/journal.pone.0039966.g001

(sputum scarce). 52% (59/113) of TB patients were sputum smear-positive.

### Urine MTB/RIF Diagnostic Accuracy

Table 2 outlines the diagnostic accuracy of sputum smear microscopy, urine MTB/RIF, LAM ELISA and LAM strip test in all culture-positive patients stratified by CD4 cell count and sputum scarce culture-positive patients only. Overall, MTB/RIF had a sensitivity of 48% (95% CI: 39–57; 54/113), equivalent to the overall sensitivities of sputum smear microscopy [52% (95% CI: 43–61; 59/113)], urine LAM ELISA [58% (95% CI: 49–67; 65/112)] and LAM strip test (grade 2 cut-point) [48% (95% CI: 39–57; 55/113)]. Urine MTB/RIF sensitivity was higher in patients with CD4 ≤ 200 cells/ml vs. CD4 ≥ 200 cells/ml [54% (95% CI: 43–65; 42/78) vs. 31% (95% CI: 17–50; 8/26), p = 0.04]. The highest urine MTB/RIF sensitivity of 61% (95% CI: 48–73; 33/54) was in CD4 ≤ 100 cells/ml. In sputum scarce, non-sputum culture-positive patients, the sensitivity of urine MTB/RIF was 40% (95% CI: 22–61; 8/20). Additionally, one of the eight urine MTB/RIF positive patients was found to have rifampicin-resistance and this was confirmed by phenotypic drug-susceptibility testing. Urine MTB/RIF sensitivity was equivalent to urine LAM ELISA and LAM strip test regardless of a patient's

ability to produce sputum or sputum smear status. In sputum smear-negative patients, the sensitivity of urine MTB/RIF was 39% (95% CI: 27–52; 21/54). The specificity of urine-based MTB/RIF was 98% (95% CI: 95–100; 61/62), which was higher than both the urine LAM ELISA and strip test [98% (95% CI: 95–100) vs. 89% (95% CI: 81–97), p = 0.03].

### The Effect of Urine Centrifugation on MTB/RIF Performance

33% (38/116) and 41% (25/61) of culture positive and negative patients respectively had sufficient archived urine available to perform MTB/RIF using both 1 ml unprocessed and 2–10 ml centrifuged and pelleted urine. The median (IQR) of urine used for pelleting was 10 (5–10) ml. Comparing paired samples, the sensitivity of urine MTB/RIF was higher using 2–10 ml centrifuged and pelleted urine than 1 ml unprocessed urine [42% (95% CI: 26–58; 16/38) vs. 8% (95% CI: 0–16; 3/38), p < 0.001]. Specificity of urine MTB/RIF was not affected by pelleting. Pelleting allowed an additional 13 cases of TB to be detected, four of which were from sputum scarce patients. Centrifugation and pelleting of 6–10 mls produced a non-significant increase in sensitivity compared with 2–5 mls [48% (95% CI: 29–67; 13/27) vs. 28% (95% CI: 0–54; 3/11), p = 0.2]. No difference was noted in

**Table 1.** Demographic, clinical and microbiological characteristics of *m.tb* culture positive and HIV-infected patients in the study population stratified by sputum smear microscopy and urine MTB/RIF results.

Patient characteristic(s)	≥1 sputum or non-sputum <i>m.tb</i> culture positive	Sputum smear positive	Sputum smear-negative or sputum scarce	Urine MTB/RIF positive, <i>m.tb</i> culture positive	Urine MTB/RIF negative, <i>m.tb</i> culture positive	P-value <sup>†</sup>
	(N = 113)	(N = 59)	(N = 54)	(N = 54)	(N = 59)	
	n (%)	n (%)	n (%)	n (%)	n (%)	
Age (years)	35	34	35	34	35	n/s
(median, IQR)	(28–38)	(27–38)	(28–38)	(28–39)	(28–38)	
Male	46 (41)	22 (37)	24 (44)	27 (50)	19 (32)	n/s
CD4 cell count <sup>#</sup>	89	109	80	56*	142*	*0.001
(median, IQR)	(45–198)	(50–215)	(40–167)	(33–134)	(59–241)	
Previous TB	32 (28)	19 (32)	13 (24)	13 (24)	19 (32)	n/s
Current Smoker	25 (22)	15 (25)	10 (19)	13 (24)	12 (20)	n/s
<b>Clinical features</b>						
Cough >2 wks	96 (85)	52 (88)	43 (80)	46 (85)	50 (85)	n/s
Night sweats	81 (72)	45 (76)	36 (67)	42 (78)	39 (66)	n/s
Weight loss	102 (90)	53 (88)	50 (93)	52 (96)*	50 (85)*	*0.003
Fever >38°C	29 (26)	19 (32)	10 (19)	16 (30)	13 (22)	n/s
<b>Clinical samples collected for TB culture</b>						
1 sputum sample	90 (80)	59 (100)*	34 (63)*	44 (82)	44 (78)	*<0.001
≥2 sputum samples	41 (36)	29 (49)*	12 (22)*	24 (44)	17 (29)	*0.003
1 non-sputum sample	75 (67)	29 (49)*	46 (85)*	35 (65)	40 (68)	*<0.001
≥2 non-sputum sample	25 (22)	7 (12)*	18 (33)*	15 (28)	10 (17)	*0.006
<b>Microbiological TB diagnosis</b>						
1 × sputum <i>m.tb</i> culture positive	87 (77)	59 (100)*	28 (52)*	42 (78)	45 (76)	*<0.001
1 × non-sputum culture <i>m.tb</i> positive	45 (40) <sup>§</sup>	14 (24)*	31 (57)*	24 (44)	21 (36)	*<0.001
1 × both sputum & non-sputum <i>m.tb</i> culture positive	19 (17)	14 (24)*	5 (9)*	12 (22)	7 (12)	*0.04

<sup>#</sup>9 patients did not have data CD4 count data.

<sup>§</sup>Includes the following *m.tb* culture positive samples: 7 blood cultures, 11 pleural fluid samples, 4 pericardial fluid samples, 3 ascitic fluid samples, 8 cerebrospinal fluid (CSF) samples, 1 hip biopsy, 2 Lymph node biopsies, 5 fine needle aspirate, 1 gastric washing, 1 faeces and 2 pus swabs.

<sup>†</sup>P-values indicate significant differences between patient groups (marked with \* to indicate comparison group) for demographic, clinical or microbiological characteristics.

doi:10.1371/journal.pone.0039966.t001

the mean (SD) MTB/RIF internal positive control C<sub>T</sub>-value for the MTB/RIF using 1 ml vs. 2–10 ml urine volumes [25.8 (1.4) vs. 25.9 (1.8), p = 0.4].

### Combined Urine LAM ELISA and MTB/RIF

Figure 2 shows the proportions of sputum scarce TB patients that were test-positive for both the urine LAM ELISA and MTB/RIF or on only one of the two tests. Both urine MTB/RIF and LAM ELISA detected six cases, while urine MTB/RIF and LAM ELISA combined detected an additional two and six cases, respectively. Table 2 shows that the combined sensitivity in sputum scarce patients of urine LAM ELISA followed by MTB/RIF was 70% (95% CI: 48–85; 14/20), compared to 40% (95% CI: 22–61; 8/20) for urine MTB/RIF (p = 0.06), 60% (95% CI: 39–78; 12/20) for LAM ELISA (p = 0.5), and 45% (95% CI: 26–66, 9/20) for LAM strip test (p = 0.1) alone. Overall, and in other relevant patient sub-groups (sputum smear-negative/sputum scarce, CD4 ≤ 200 or ≤ 100 cells/ml), the combined sensitivity of urine LAM ELISA and MTB/RIF was better than urine

MTB/RIF and LAM strip tests alone, but not urine LAM ELISA alone.

### Predictors of Urine MTB/RIF Test Positivity

Table 3 presents univariate associations of MTB/RIF positivity in HIV-infected culture-positive patients stratified by sputum smear status. CD4 ≤ 50 cells/ml, protein/creatinine ratio > 0.03 g/l and urine LAM ELISA positivity were strong predictors of urine MTB/RIF positivity. Additionally, patients with a glomerular filtration rate (GFR) between 30–60 ml/min were more likely to have a positive MTB/RIF [OR (95% CI): 3.0 (1.0–8.3), p = 0.04]. Previous TB, smoking status, any TB symptom, admission vital signs and urine dipstick abnormalities were not associated with MTB/RIF positivity. In sputum smear/negative and sputum scarce patients, only a CD4 ≤ 50 and a protein/creatinine ratio > 0.03 g/l were strong predictors of urine MTB/RIF positivity.

**Table 2.** Diagnostic accuracy of sputum smear microscopy, urinary MTB/RIF, TB LAM ELISA, LAM strip test (grade 2 cut-point) and clinically relevant combinations thereof in any sputum/non-sputum *m.tb* culture positive patients overall, in sputum scarce patients only, and stratified by CD4 cell count<sup>†</sup>.

Diagnostic test(s)	All <i>m.tb</i> culture positive	Only sputum-scarce non-sputum <i>m.tb</i> culture positive	HIV-infected patients with CD4 count >200 cells/ml	HIV-infected patients with CD4 count ≤200 cells/ml	Random sample of <i>m.tb</i> culture negative patients <sup>§</sup>
	(N = 113)	(N = 20)	(N = 26)	(N = 78)	(N = 62)
	Sensitivity (%)	Sensitivity (%)	Sensitivity (%)	Sensitivity (%)	Specificity (%)
	(95% CI)	(95% CI)	(95% CI)	(95% CI)	(95% CI)
	n/N	n/N	n/N	n/N	n/N
<b>Sputum smear microscopy</b>	52 <sup>#3</sup> (43–61) 59/113	N/A	58 <sup>#6 #7</sup> (39–75) 15/26	50 <sup>#8</sup> (39–61) 39/78	100 (94–100) 62/62
<b>Urine MTB/RIF</b>	48 <sup>#1</sup> (39–57) 54/113	40 <sup>#5</sup> (22–61) 8/20	31 <sup>#7 *1</sup> (17–50) 8/26	54 <sup>#9 *1</sup> (43–65) 42/78	98 <sup>#11</sup> (95–100) 61/62
<b>Urine LAM ELISA</b>	58 <sup>#4</sup> (49–67) 65/112	60 (39–78) 12/20	27 <sup>*2</sup> (14–46) 7/26	69 <sup>#8 *2</sup> (58–78) 53/77	89 <sup>#11</sup> (81–97) 55/62
<b>Urine LAM strip test (grade 2 cut-point)</b>	48 <sup>#2</sup> (39–57) 55/113	45 (26–66) 9/20	27 <sup>#6 *3</sup> (14–46) 7/26	56 <sup>#10 *3</sup> (45–67) 44/78	85 (77–94) 53/62
<b>Urine LAM ELISA followed by urine MTB/RIF (performed if LAM ELISA negative)</b>	68 <sup>#1 #2</sup> (60–77) 77/113	70 <sup>#5</sup> (48–85) 14/20	38 <sup>*4</sup> (20–57) 10/26	79 <sup>#9 #10 *4</sup> (71–88) 62/78	89 (81–97) 55/62
<b>Urine LAM ELISA combined with smear microscopy</b>	74 <sup>#3 #4</sup> (65–82) 83/113	N/A	58 <sup>*5</sup> (39–77) 15/26	80 <sup>*5</sup> (71–88) 62/78	89 (81–97) 55/62

<sup>#</sup>Indicates p<0.05 for a comparison of the sensitivity between different tests (e.g. urine MTB/RIF vs. LAM strip test) or combinations thereof; specific p-value: <sup>#1</sup>p = 0.002; <sup>#2</sup>p = 0.003; <sup>#3</sup>p = 0.001; <sup>#4</sup>p = 0.01; <sup>#5</sup>p = 0.06; <sup>#6</sup>p = 0.02; <sup>#7</sup>p = 0.05; <sup>#8</sup>p = 0.02; <sup>#9</sup>p < 0.001; <sup>#10</sup>p = 0.002; <sup>#11</sup>p = 0.03.

<sup>\*</sup>Indicates p<0.05 for a comparison of differences in sensitivity between CD4>200 and ≤200 groups for a specific test or combinations thereof; specific p-values: <sup>\*1</sup>p = 0.04; <sup>\*2</sup>p < 0.001; <sup>\*3</sup>p = 0.009; <sup>\*4</sup>p < 0.001; <sup>\*5</sup>p = 0.03; non-significant p-values not shown.

<sup>†</sup>9 culture positive patients with no available CD4 cell count results.

<sup>§</sup>Reference standard of culture was used which does not account for persons with culture-negative, clinical TB.

doi:10.1371/journal.pone.0039966.t002

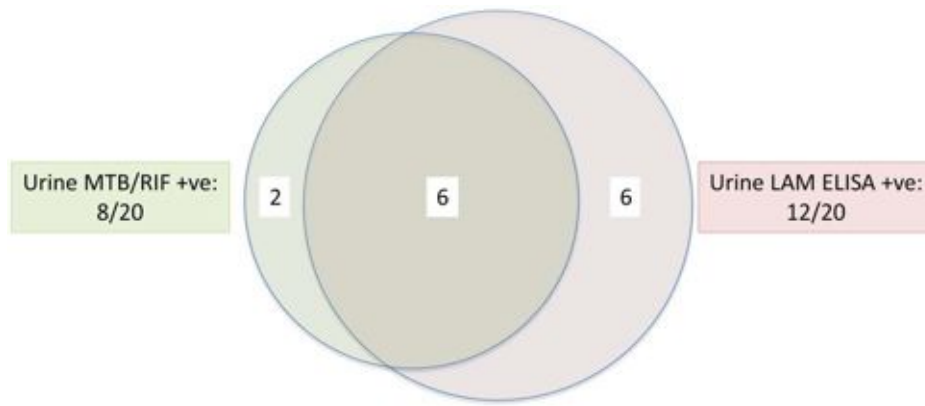
### Urine MTB/RIF C<sub>T</sub> Values and Relationship to Other Bacillary Burden Markers

In urine MTB/RIF positive patients, the mean (sd) C<sub>T</sub> -value was 21.3 (13.8). Two-way scatter plots in figure 3 explore correlations between urine C<sub>T</sub> -values and other markers of bacillary burden. No correlations between urine MTB/RIF-generated C<sub>T</sub> -values and mean urine LAM concentration [spearman rho: 0.17, p = 0.25], LAM strip test grade [spearman rho: 0.11, p = 0.4], sputum liquid culture time-to-positivity (TTP) [spearman rho: 0.23, p = 0.15] and/or sputum smear microscopy grade (in smear positive patients) [spearman rho: 0.04, p = 0.8] were found. No correlation between urine C<sub>T</sub> -values and urine culture TTP was possible as only 1/113 urine cultures was positive. A weak, but significant inverse relationship between urine C<sub>T</sub> -value and liquid culture TTP [spearman rho: 0.3, p = 0.03] was noted. Additionally, urine MTB/RIF-generated C<sub>T</sub>-values

showed no significant correlation with CD4 T-cell count [spearman rho: 0.27, p = 0.06].

### Discussion

Sputum scarce, smear-negative and EPTB constitute a major burden of undiagnosed TB in HIV co-infected hospitalised patients [5]. Misdiagnosis and diagnostic delays lengthen hospital stay and delay rapid treatment initiation, likely worsening TB-related morbidity and mortality [8]. Urine-based diagnostics may be particularly useful in these patients. The key findings of this preliminary study are: i) urine MTB/RIF can detect TB, and rifampicin resistance where applicable, in hospitalised HIV-infected patients who cannot produce a sputum sample (the assay detected 8 cases who would have been missed by conventional sputum-based diagnostics i.e. 7% of culture positive TB cases); ii) centrifugation and pelleting of 2–10 mls of urine significantly



**Figure 2. Venn diagram showing the proportions of patients diagnosed by urine MTB/RIF and/or urine LAM ELISA in sputum scarce *m.tb* culture positive patients.** 6 patients were both urine MTB/RIF and LAM ELISA positive; 6 patients were urine LAM ELISA positive, but MTB/RIF negative, while 2 patients were urine MTB/RIF positive, LAM ELISA negative.  
doi:10.1371/journal.pone.0039966.g002

improved diagnostic yield vs. 1 ml unprocessed urine without impacting the MTB/RIF indeterminate rate (almost a third of patients diagnosed exclusively using centrifuged urine were sputum scarce); iii) in sputum scarce, and other patient sub-groups, the combination of urine LAM ELISA with urine MTB/RIF improved sensitivity compared with urine MTB/RIF or LAM strip tests alone; and iv) urine MTB/RIF positivity was strongly associated with advanced immunosuppression and proteinuria (suggesting a 'leaky' renal filtration mechanism).

Approximately 10–20% of HIV-infected patients with suspected TB are unable to produce sputum for sputum-based diagnostics [4,17]. An additional 50–70% will have smear-negative active TB [18]. Sputum induction by ultrasonic nebulisation is an important

sampling adjunct but availability is very limited in resource-poor settings, infection control is problematic, and sputum is still unobtainable in ~10% of patients [17]. Thus, urine-based diagnostics have the potential to offer incremental diagnostic yield in this context. Urinary LAM, both as an ELISA-kit and recently as a lateral flow strip test, is the most extensively evaluated urine-based TB diagnostic [9,13,19]. It is most useful as a rule-in test in HIV-infected persons with advanced immunosuppression [4,20,21]. However, concerns about context-specific test specificity, optimal cut-point selection (for the LAM strip test), and lack of a drug-susceptibility read-out remain [4]. In this study we show the overlapping, yet non-redundant, performance of urinary LAM and MTB/RIF, and the improved overall diagnostic sensitivity of using the tests in combination. There is limited information about the performance of MTB/RIF in combination with adjunct diagnostics [22].

There are limited published data about the performance of MTB/RIF using urine samples. Hilleman *et al.*, in a selected laboratory cohort interrogating extra-pulmonary samples found that MTB/RIF sensitivity was 100% in 6 culture-positive urine samples with unknown HIV status [23], while Lawn *et al.*, in HIV-infected out-patients pre-ARV initiation, report the overall sensitivity of urine MTB/RIF to be 19% [24]. By contrast, our study focused on sputum scarce, hospitalised HIV-infected patients where diagnosis is often delayed and challenging. It is likely that HIV-infected patients' with more advanced immunosuppression accounted for the higher urine MTB/RIF sensitivity found in our study.

We found a strong association between declining CD4 cell count, LAM in the urine, proteinuria, and increasing urine MTB/RIF positivity. This may reflect renal TB as part of disseminated TB, increased bacillary burden in those with the most advanced immunosuppression, a 'leaky' filtration mechanism or a combination of these. That a minority of the MTB/RIF-positive samples were urine culture positive may simply reflect sampling error, the limited volume sent for culture, or that there are several mechanisms driving urine test positivity. Further molecular biological and pathological studies are required to shed more light on the underlying mechanisms.

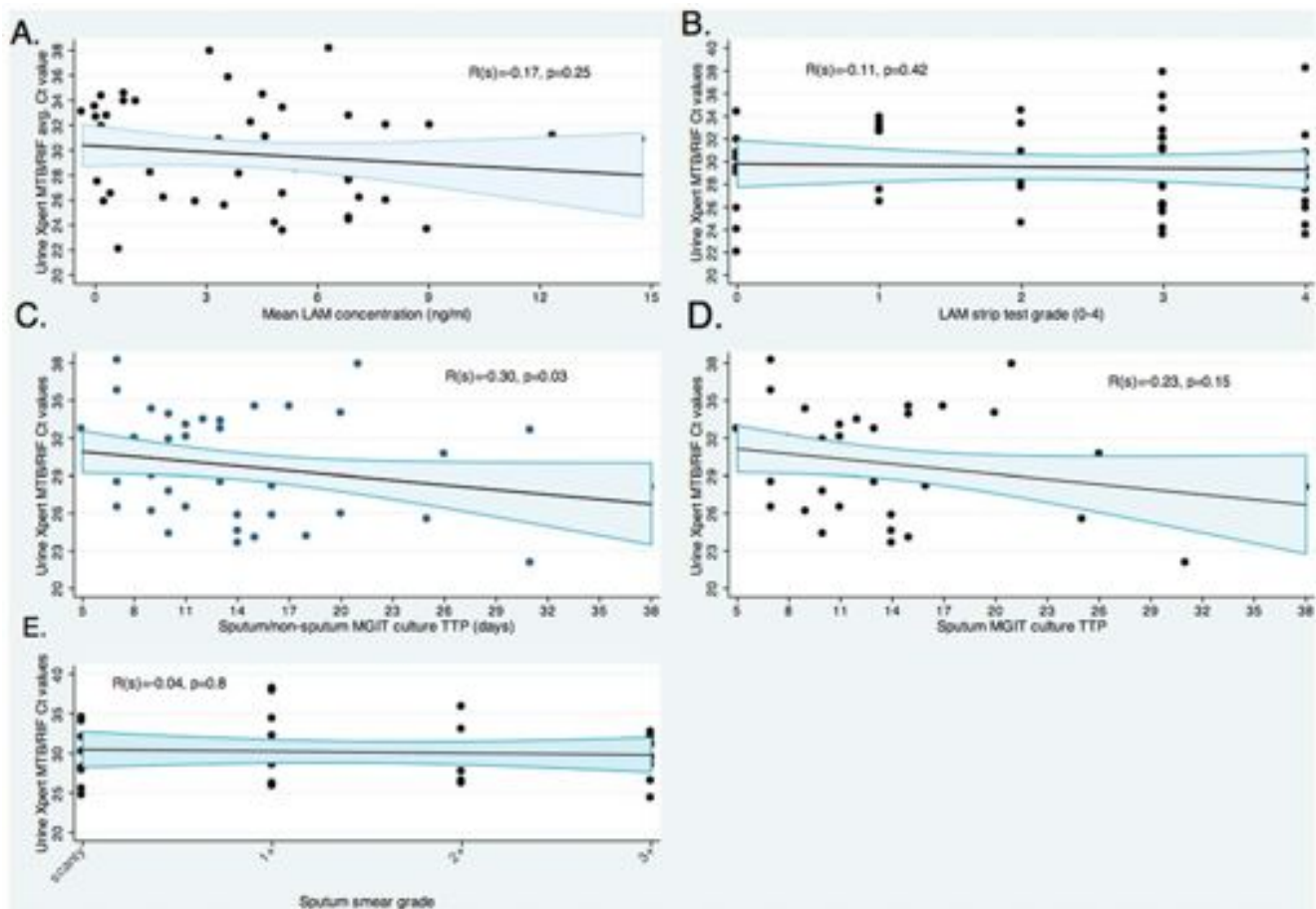
The sensitivity of urine MTB/RIF was markedly improved by the centrifugation and pelleting of ~ 2–10 mls urine. Indeed, the concentration of a number of biological samples, such as cerebrospinal and pleural fluid, has improved the performance of traditional TB diagnostics [25]. However, concentration is also

**Table 3. Associates of MTB/RIF positivity in HIV-infected *m.tb* culture positive patients stratified by smear status.**

Patient characteristic (s)	Odds ratio (95% CI)	P-value
<b>All patients (sputum smear positive and negative)</b>		
CD4≤200	2.6 (1.0–6.7)	0.05
CD4≤100	3.1 (1.4–6.8)	0.006
CD4≤50	5.3 (2.0–13.9)	0.001
Protein/creatinine ratio >0.03 g/l	6.2 (2.3–17.1)	<0.001
Urea≥7.1 mmol/l	1.6 (0.7–3.9)	0.3
GFR 30–60 ml/min*	3.0 (1.0–8.3)	0.04
Urine LAM ELISA	5.0 (2.2–11.4)	<0.001
<b>Sputum smear-negative/scarce patients only</b>		
Age (years)	1.08 (1.00–1.16)	0.04
CD4≤200	3.9 (0.7–20.2)	0.11
CD4≤100	5.2 (1.4–19.4)	0.01
CD4≤50	9.8 (2.4–38.8)	0.001
Protein/creatinine ratio >0.03 g/l	5.3 (1.3–21.8)	0.02
Urea≥7.1 mmol/l	1.3 (0.3–4.8)	0.3
GFR 30–60 ml/min*	2.4 (0.4–14.0)	0.3
Urine LAM ELISA	2.6 (0.9–8.2)	0.1

\*Glomerular filtration rate calculated using the modified Cochrane-Gault equation.

doi:10.1371/journal.pone.0039966.t003



**Figure 3. Urine-based MTB/RIF-generated  $C_T$  values correlates with other markers of bacillary burden.** In *m.tb* culture and urine MTB/RIF positive patients, MTB/RIF-generated  $C_T$  values were correlated with (A) urine LAM ELISA concentration, (B) urine LAM strip grading, (C) either sputum or non-sputum liquid culture time-to-positivity (TTP), (D) only sputum liquid culture time-to-positivity, and (E) smear-grade in smear positive patients.

doi:10.1371/journal.pone.0039966.g003

believed to increase PCR inhibition and potentially the rate of indeterminate test results. With only a single indeterminate test result (on a 1 ml urine sample), and no change in the mean internal positive control  $C_T$ -value in centrifuged samples, we feel that 10 mls of centrifuged and pelleted urine is optimal when using MTB/RIF. The incremental yield of using volumes greater than 10 mls will require further evaluation.

Sputum-based MTB/RIF-generated  $C_T$ -values have been shown to strongly correlate with other markers of bacillary load, such as liquid culture TTP and smear grade [15,26,27]. However, we could show no such correlation with urine (either with urine or sputum TTP and smear grade). This lack of correlation likely reflects limited sample numbers, and the differential relationship between urine bacillary load, renal abnormalities (renal TB, glomerular dysfunction etc.) and total body bacillary load.

This preliminary study has important limitations. No sputum-based MTB/RIF was performed in our study, thus we could not compare performance between urine and sputum. However, sputum samples were not stored in the parent study [4] and this caveat is redundant in sputum scarce patients, which is the very subgroup targeted by our study. Archived, frozen urine samples were used for all TB diagnostic tests. This may have affected the diagnostic performance of urine culture, but prior studies suggest

no impact on the performance of MTB/RIF [28] or LAM [19] in this context.

In conclusion, this preliminary study indicates that urinary MTB/RIF may aid the rapid diagnosis of TB in sputum scarce HIV-infected patients with advanced immunosuppression. Moreover, urine centrifugation significantly improves sensitivity. Used alone, or in combination with urine LAM, urine-based MTB/RIF may potentially offer ~70% of HIV co-infected persons a TB diagnosis with 24 hours of hospital admission. Finally, as demonstrated in this study, the ability of MTB/RIF to offer rifampicin drug susceptibility testing is an important advantage over urine LAM. Further prospective studies in larger cohorts using standardised TB diagnostic work-up are required to clarify these findings.

### Acknowledgments

We thank the research nursing staff, including Sr Nicola Kelly and Faranaaz Embrahim, who were involved in the recruitment of patients and collection of urine samples. The support of the clinical and nursing staff of the general medical wards of Groote Schuur, Victoria, GF Jooste and New Somerset hospitals was greatly appreciated.

## Author Contributions

Conceived and designed the experiments: JP GT TM UG KD. Performed the experiments: TM UG GT. Analyzed the data: JP GT KD. Wrote the paper: JP GT TM UG KD.

## References

- WHO (2011) WHO Report 2011: Global tuberculosis control. Geneva: World Health Organisation. ISBN 978 92 4 156438 0 ISBN 978 92 4 156438 0.
- Reid MJ, Shah NS (2009) Approaches to tuberculosis screening and diagnosis in people with HIV in resource-limited settings. *Lancet Infect Dis* 9: 173–184.
- Jones BE, Young SM, Antoniskis D, Davidson PT, Kramer F, et al. (1993) Relationship of the manifestations of tuberculosis to CD4 cell counts in patients with human immunodeficiency virus infection. *The American review of respiratory disease* 148: 1292–1297.
- Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, et al. (2012) Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*.
- Cox JA, Lukande RL, Lucas S, Nelson AM, Van Marck E, et al. (2010) Autopsy causes of death in HIV-positive individuals in sub-Saharan Africa and correlation with clinical diagnoses. *AIDS Rev* 12: 183–194.
- Rana FS, Hawken MP, Mwachari C, Bhatt SM, Abdullah F, et al. (2000) Autopsy study of HIV-1-positive and HIV-1-negative adult medical patients in Nairobi, Kenya. *J Acquir Immune Defic Syndr* 24: 23–29.
- Cohen T, Murray M, Wallengren K, Alvarez GG, Samuel EY, et al. (2010) The prevalence and drug sensitivity of tuberculosis among patients dying in hospital in KwaZulu-Natal, South Africa: a postmortem study. *PLoS Med* 7: e1000296.
- Holtz TH, Kabera G, Mthiyane T, Zingoni T, Nadesan S, et al. (2011) Use of a WHO-recommended algorithm to reduce mortality in seriously ill patients with HIV infection and smear-negative pulmonary tuberculosis in South Africa: an observational cohort study. *The Lancet infectious diseases* 11: 533–540.
- Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, et al. (2010) Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. *Curr Opin Pulm Med* 16: 262–270.
- WHO (2010) Roadmap for Rolling Out Xpert MTB/RIF for Rapid Diagnosis of TB and MDR-TB World Health Organization, Geneva, Switzerland.
- WHO (2011) Rapid Implementation of the Xpert MTB/RIF diagnostic test. Publication number WHO/HTM/TB/2011.2. World Health Organization, Geneva, Switzerland.
- Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, et al. (2011) Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the Xpert MTB/RIF test for diagnosis of tuberculosis and multidrug resistance: a multicentre implementation study. *Lancet* 377: 1495–1505.
- Flores LL, Steingart KR, Dendukuri N, Schiller I, Minion J, et al. (2011) Systematic review and meta-analysis of antigen detection tests for the diagnosis of tuberculosis. *Clinical and vaccine immunology : CVI* 18: 1616–1627.
- Dhedra K, Van-Zyl Smit RN, Sechi LA, Badri M, Meldau R, et al. (2009) Clinical diagnostic utility of IP-10 and LAM antigen levels for the diagnosis of tuberculous pleural effusions in a high burden setting. *PLoS One* 4: e4689.
- Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, et al. (2011) Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. *American journal of respiratory and critical care medicine* 184: 132–140.
- Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. (2003) The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *The Standards for Reporting of Diagnostic Accuracy Group. Croat Med J* 44: 639–650.
- Gonzalez-Angulo Y, Wiysonge CS, Geldenhuys H, Hanekom W, Mahomed H, et al. (2011) Sputum induction for the diagnosis of pulmonary tuberculosis: a systematic review and meta-analysis. *European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology*.
- Lawn SD, Brooks SV, Kranzer K, Nicol MP, Whitelaw A, et al. (2011) Screening for HIV-associated tuberculosis and rifampicin resistance before antiretroviral therapy using the Xpert MTB/RIF assay: a prospective study. *PLoS Med* 8: e1001067.
- Minion J, Leung E, Talbot E, Dheda K, Pai M, et al. (2011) Diagnosing tuberculosis with urine lipoarabinomannan: systematic review and meta-analysis. *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology* 38: 1398–1405.
- Shah M, Variava E, Holmes CB, Coppin A, Golub JE, et al. (2009) Diagnostic accuracy of a urine lipoarabinomannan test for tuberculosis in hospitalized patients in a High HIV prevalence setting. *Journal of acquired immune deficiency syndromes* 52: 145–151.
- Lawn SD, Kerkhoff AD, Vogt M, Wood R (2011) 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*.
- Theron G, Pooran A, Peter J, van Zyl-Smit R, Mishra HK, et al. (2011) Do adjunct TB tests, when combined with Xpert MTB/RIF, improve accuracy and the cost of diagnosis in a resource-poor setting? *The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology*.
- Hillemann D, Rusch-Gerdes S, Boehme C, Richter E (2011) Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. *J Clin Microbiol* 49: 1202–1205.
- Lawn SD, Kerkhoff AD, Vogt M, Wood R (2012) High diagnostic yield of tuberculosis from screening urine samples from HIV-infected patients with advanced immunodeficiency using the Xpert MTB/RIF assay. *Journal of Acquired Immune Deficiency Syndromes*.
- Stewart SM (1953) The bacteriological diagnosis of tuberculous meningitis. *Journal of clinical pathology* 6: 241–242.
- Blakemore R, Nabeta P, Davidow AL, Vadwai V, Tahiri R, et al. (2011) A multisite assessment of the quantitative capabilities of the Xpert MTB/RIF assay. *American journal of respiratory and critical care medicine* 184: 1076–1084.
- Theron G, Pinto L, Peter J, Mishra HK, van Zyl-Smit R, et al. (2012) The use of an automated quantitative polymerase chain reaction (Xpert MTB/RIF) to predict the sputum smear status of tuberculosis patients. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 54: 384–388.
- Moure R, Munoz L, Torres M, Santin M, Martin R, et al. (2011) Rapid detection of *Mycobacterium tuberculosis* complex and rifampin resistance in smear-negative clinical samples by use of an integrated real-time PCR method. *J Clin Microbiol* 49: 1137–1139.

**Chapter 8.**

**The relationship between urine LAM concentration and  
compartment-specific TB bacillary load in hospitalised HIV-infected  
patients with advanced immunosuppression**

Peter, J. G., Theron, G. and Dheda, K.

Published in *South African Respiratory Journal* 2012 Volume 18 Issue 2: 43-46

University of Cape Town

# The relationship between urine LAM concentration and compartment-specific TB bacillary load in hospitalised HIV-infected patients with advanced immunosuppression

Jonny Peter<sup>1</sup>, Grant Theron<sup>1</sup>, Keertan Dheda<sup>1,2</sup>

<sup>1</sup>Lung Infection and Immunity Unit, Division of Pulmonology and UCT Lung Institute, Department of Medicine, Faculty of Health Sciences, University of Cape Town, Cape Town, South Africa.

<sup>2</sup>Institute of Infectious Diseases and Molecular Medicine, Faculty of Health Sciences, UCT, Cape Town, South Africa

## Abstract

### Background:

Urinary lipoarabinomannan (LAM) detection, using either an ELISA-kit or the novel point-of-care Determine®-TB lateral flow strip test, offers potential for rapid TB diagnosis in HIV-infected patients with advanced immunosuppression. High bacillary burden, as measured by smear grade or liquid culture time-to-detection, correlates with poor outcomes in smear-positive patients. We investigated urine LAM concentration as a potential useful marker of total bacillary burden in hospitalised HIV-infected patients with advanced immunosuppression.

### Methods:

In a well-characterised cohort of 281 HIV-infected hospitalised patients with advanced immunosuppression the relationship between traditional and novel compartment-specific markers of bacillary burden (smear positivity and grade [lung]; time-to-detection on sputum and non-sputum liquid culture [lung] and; urine MTB/RIF-generated CT values [kidney]) and urine LAM concentrations were determined.

### Results:

Urine LAM concentrations were higher amongst smear-positive compared to smear-negative TB patients ( $p=0.03$ ), and patients with  $CD4 < 100$  cells/ml compared to  $> 100$  cells/ml ( $p=0.02$ ). No correlation was found between LAM concentration and the time-to-detection on liquid culture, or urine MTB/RIF-generated CT-values. The strongest predictors of urine LAM concentrations were urine MTB/RIF positivity ( $p=0.009$ ), a decreasing glomerular filtration rate ( $p=0.02$ ) and an elevated protein/creatinine ratio ( $p=0.004$ ).

### Conclusions:

Higher urine LAM concentrations are found in persons with a high sputum bacillary burden and those with more advanced immunosuppression. However, the presence of renal TB (urine MTB/RIF positivity) and abnormalities of renal filtration/function also significantly impact the urine LAM concentration.

**Key words:** TB HIV co-infection, urine LAM, bacillary burden

## Introduction

In tuberculous sputum smear-positive patients an increased bacillary burden, as measured by the time-to-detection on liquid culture, has been found to be an important predictor of delayed two-month culture conversion and poor outcomes.<sup>1,2</sup> However, in HIV-infected patients with smear-negative and disseminated forms of TB disease, demonstrating the impact of bacillary burden is more difficult because:

- i) obtaining adequate samples is a challenge, especially from extra-pulmonary sites, and
- ii) compartment specific bacillary load may not accurately reflect total bacillary burden.

The need for biomarkers that provide a good measure of total bacillary burden in these patients is needed.

Lipoarabinomannan (LAM) is a 17.5 kDa glycolipid component of the mycobacterial cell wall. LAM is heat stable, filtered by the kidneys and can be detected in urine.<sup>3</sup> Urine LAM has been extensively studied and offers important potential utility as a diagnostic adjunct for TB detection in HIV-infected patients with advanced immunosuppression.<sup>4,5</sup> Initially developed as an ELISA-kit (now marketed as the TB LAM ELISA, Alere, USA), it has now been superseded by the point-of-care Determine®-TB lateral flow strip test.<sup>4</sup> This affordable (<US\$3.5) urine LAM strip test could offer bedside, laboratory-free diagnosis from just 60µl of urine in 25 minutes.<sup>4</sup> In addition to diagnosis, both the original TB LAM ELISA and novel LAM strip test can offer quantification and semi-quantification of urinary LAM respectively.

Urine LAM concentration has been shown to increase with smear grade and be higher in disseminated (sputum and blood sample culture positive) vs isolated pulmonary (sputum culture positive only) or extra-pulmonary (non-sputum culture positive only) TB disease.<sup>6</sup> However, the relationship of LAM concentration to more accurate measures of bacillary burden such as culture time-to-detection or MTB/RIF-generated cycle threshold values is unclear.<sup>7</sup> Furthermore, few data are available on predictors of urine LAM concentration. Thus, we investigated the relationship between urine LAM concentrations and other traditional and novel measurements of bacillary load, and identified the strongest determinates of urine LAM concentration.

## Materials and Methods

### Study population

The study population, TB diagnostic work-up as well as the clinical, radiological and laboratory parameters collected for patients in this cohort study has been previously described.<sup>5</sup> Fluorescence smear microscopy was performed on NALC/NaOH processed and concentrated samples followed by liquid culture (BACTEC MGIT 960, BD Diagnostics, USA); all clinical specimens were processed by the National Health Laboratory services. Smear grading methodology has been previously described<sup>8</sup> and the automated MGIT 960 provided time-to-detection for liquid culture. All patients provided written informed consent and the study was approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee.

### Urine testing and methodology

A spot urine sample (10-30ml) was collected, frozen and stored at -20°C for later batched testing. TB LAM ELISA was performed on thawed urine according to the manufacturers' instructions as previously described.<sup>9</sup> Urine LAM concentrations were determined from the optical density ELISA measurements using standard curves. Additional urine testing included the Xpert MTB/RIF assay and the measurement of protein-creatinine ratios. MTB/RIF-generated cycle threshold (CT) values correspond to the number of PCR cycles required before *M. tuberculosis* genetic material is detected; MTB/RIF-generated CT values are inversely correlated with the amount of intact *M. tuberculosis* organisms (viable or non-viable) in the starting sample material.<sup>10</sup>

### Statistical analysis

Only patients with a positive *M. tuberculosis* liquid culture were used in the analysis. Wilcoxon rank-sum and Kruskal Wallis tests, as appropriate, were used to compare the median (interquartile range) urine LAM concentration between groups. Spearman correlation (Rs) was used to evaluate relationship between urine LAM concentrations, liquid culture time-to-detection and MTB/RIF-generated CT values. Linear regression was used to determine the strongest predictors of increasing urine LAM concentrations. STATA IC, version 10 (Stata Corp, Texas, USA) was used for all statistical analyses.

## Results

### Urine LAM concentration and other markers of bacillary load

The median (IQR) urine LAM concentration was 0.64 (0.07-5.05) ng/ml in *M. tuberculosis* culture positive patients. **Figure 1** shows box and whisker plots of urine LAM concentration by sputum smear grade and CD4 cell count strata. Median (IQR) LAM concentration was higher in smear positive vs. smear negative patients [2.93 (0.11-6.48) vs. 0.19 (0.05-0.78),  $p=0.003$ ]. In patients with CD4  $\leq 100$ , 101-200 and  $>200$  cells/ml the median (IQR) LAM concentrations were 3.23 (0.27-5.64), 0.43 (0.07-3.27) and 0.09 (0-4.4) ( $p=0.01$ ). In all patients with  $\geq 1$  sputum and non-sputum sample, the median (IQR) LAM concentration was 2.09 (0.12-5.05) in patients with both a sputum and non-sputum culture positivity (disseminated), 0.33 (0.06-5.05) in patients with only a

non-sputum culture positive (extra-pulmonary) and 0.21 (0.65-3.34) in patients with only a sputum culture positive (pulmonary) ( $p=0.4$ ). **Figure 2** shows scatter plots and linear regression lines for urine LAM concentration (ng/ml) and i) sputum/non-sputum liquid culture time-to-detection (days) and

ii) urine MTB/RIF-generated CT-values.

No correlation was found between urine LAM concentration and either liquid culture time-to-detection ( $r_s = -0.3$ ,  $p=0.1$ ) or urine MTB/RIF generated CT values ( $r_s = -0.1$ ,  $p=0.5$ ); no correlations were found even when stratified by type and/or number of culture positive samples or CD4 cell count.

### Predictors of urine LAM concentration

**Table 1** shows the results of the multivariate analysis for predictors of increasing LAM concentration. A positive urine MTB/RIF assay ( $p=0.009$ ), increasing protein/creatinine ratio ( $p=0.004$ ) and a decreasing glomerular filtration rate ( $p=0.02$ ) were the only significant predictors of increasing urine LAM concentration. CD4 cell count was associated with urine LAM concentration in the univariate but not multivariate analysis. No other clinical, radiological or laboratory predictors were found to significantly predict urine LAM concentrations.

## Discussion

Increased baseline sputum bacillary load in smear-positive patients have been shown to be an important predictor of poor treatment outcomes.<sup>1,2</sup> Disseminated TB is increased in HIV-infected patients with advanced immunosuppression, has a high mortality,<sup>11</sup> and it is biologically plausible that total bacillary burden may predict outcome in these patients.

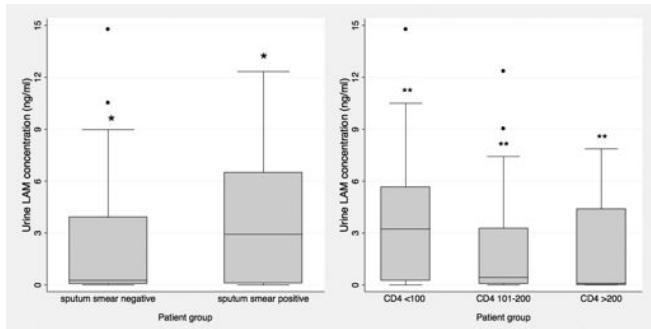
**Table 1:**  
Multivariate linear regression of clinical variables predicting urinary LAM level<sup>†</sup>

Clinical variable	$\beta$ -coefficient	P-value
Age (years)	0.004	0.9
Weight (kgs)	-0.011	0.6
Temperature (°C)	-0.039	0.8
CD4 cell count* (cells/ml)	-0.23	0.2
Urea (mmol/l)*	0.67	0.1
<b>GFR (ml/min)<sup>§</sup></b>	<b>0.018</b>	<b>0.02</b>
<b>Urine protein/creatinine ratio* (mg/mmol)</b>	<b>1.97</b>	<b>0.004</b>
<b>Urine MTB/RIF positive</b>	<b>1.21</b>	<b>0.009</b>
Time to sputum culture positivity (days)*	0.29	0.5

<sup>†</sup> Clinical variables included in the multivariate linear regression were significantly associated with urine LAM concentration in univariate linear regression or considered as potentially important confounding variables (eg. weight, age)

\* These variables were not normally distributed and were therefore the log values were used for the linear regression analysis.

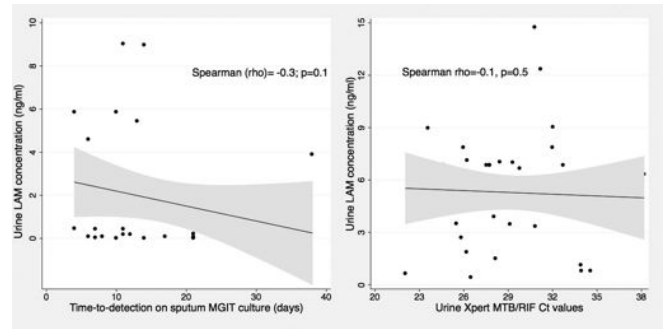
<sup>§</sup> GFR: Glomerular filtration rate; calculated using the modified Cochrane-Gault equation



**Figure 1**

Median (IQR) urine LAM concentrations (ng/ml) for i) smear positive vs. negative patients, and ii) patients with CD4 $\leq$ 100, 101-200, and >200 cells/ml.

\*p=0.003; \*\*p=0.01



**Figure 2**

Correlation of urine LAM concentration (ng/ml) with liquid culture time-to-detection in sputum/non-sputum samples and urine MTB/RIF-generated cycle threshold values.

However, traditional markers of mycobacterial burden from a single body compartment e.g. sputum may provide a limited or misleading picture. A more comprehensive picture may be obtained by sampling multiple body compartments e.g. blood, sputum and pleural fluid, but this approach is limited by the challenges of sampling in patients with advanced immunosuppression and the impact of poor sample quality. Furthermore, the relationship between compartment-specific bacillary load e.g. sputum and total bacillary burden is poorly understood. A user-friendly biomarker, such as urine LAM, that could potential provide a single measure reflective of total bacillary burden is appealing.

To date, the concentration of urine LAM, released from actively metabolising and degrading mycobacteria at any body site, has been shown to increase with an increasing number of culture positive body sites, sputum smear grade and a declining CD4 cell count.<sup>6</sup> Urine LAM positivity offers prognostic utility and has been shown to be associated with increased illness severity and mortality.<sup>6,12</sup> Furthermore, urine LAM detects a vulnerable group of severely ill hospitalised patients not identified by basic clinical and radiological screening (Peter and Dheda submitted for publication).

This study demonstrates that, although urine LAM concentration does to some extent reflect bacillary load (as measured by sputum smear positivity), no direct linear correlation could be found between urine LAM concentration and the time-to-detection in sputum or non-sputum liquid cultures. The failure to demonstrate a relationship may be confounded by sampling error (inability to acquire representative samples and poor sample quality) or the lack of multiple body compartment sampling, but more likely indicates

i) the complex relationship between individual body-site and total bacillary load, and

ii) the presence of additional and more robust determinates of urine LAM concentration.

In fact, our study findings suggest that the presence of *M.tuberculosis* within the urinary tract, as indicted by MTB/RIF positivity, together with abnormalities in both renal function and glomerular membrane permeability, are the strongest predictors of urine LAM concentration. Renal function and glomerular dysfunction may be occurring secondary to the presence of renal TB itself (as suggested by a strong correlation between urine MTB/RIF positivity and elevated protein/creatinine ratio<sup>13</sup>) or HIV nephropathy. Further research studies with standardised multiple-compartment sampling and renal biopsy data are now needed

in HIV-infected patients with advanced immunosuppression to confirm these findings, and determine novel methods to better estimate total bacillary burden in these vulnerable patients.

### Funding

This work was supported by an EU-FP7 award (TBSusgent). JP is supported by a Discovery Foundation Fellowship, the Fogarty International Clinical Research Scholars/Fellows Support Centre National Institutes of Health grant R24TW007988, SATBAT (JP) and the EDCTP (TB-NEAT). GT are supported by the EDCTP (TB-NEAT), the South African National Research Foundation (GT) and the Claude Leon Foundation (GT). KD is supported by the EDCTP (TB-NEAT, TESA), the SA DST and the SA NRF (SARChI).

### Statement of interest

The authors declare no conflicts of interest. The urine LAM ELISA kits were donated by the Alere. Alere had no role in the study or decision to publish.

### References

- Hesseling AC, Walzl G, Enarson DA, Carroll NM, Duncan K, Lukey PT, *et al.* Baseline sputum time to detection predicts month two culture conversion and relapse in non-HIV-infected patients. *International journal of tuberculosis and lung disease.* 2010 May; **14**(5):560-70.
- Visser ME, Stead MC, Walzl G, Warren R, Schomaker M, Grewal HM, *et al.* Baseline predictors of sputum culture conversion in pulmonary tuberculosis: importance of cavities, smoking, time to detection and W-Beijing genotype. *PLoS One.* 2012; **7**(1):e29588.
- Hamasur B, Bruchfeld J, Haile M, Pawlowski A, Bjorvatn B, Kallenius G, *et al.* Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine. *J Microbiol Methods.* 2001 May; **45**(1):41-52.
- Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, Dheda K. Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. *Curr Opin Pulm Med.* 2010 May; **16**(3):262-70.
- Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, Kraus S, *et al.* Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients. *European respiratory journal.* 2012 Epub Feb 23.
- Shah M, Martinson NA, Chaisson RE, Martin DJ, Variava E, Dorman SE. Quantitative analysis of a urine-based assay for detection of lipoarabinomannan in patients with tuberculosis. *J*

- Clin Microbiol.* 2010 Aug; **48**(8):2972-4.
7. van Zyl-Smit RN, Binder A, Meldau R, Mishra H, Semple PL, Theron G, et al. Comparison of quantitative techniques including Xpert MTB/RIF to evaluate mycobacterial burden. *PLoS One.* 2011; **6**(12):e28815.
  8. Whitelaw A, Peter J, Sohn H, Viljoen D, Theron G, Badri M, et al. Comparative cost and performance of light-emitting diode microscopy in HIV-tuberculosis-co-infected patients. *European respiratory journal.* 2011 Dec; **38**(6):1393-7.
  9. Dheda K, Davids V, Lenders L, Roberts T, Meldau R, Ling D, et al. Clinical utility of a commercial LAM-ELISA assay for TB diagnosis in HIV-infected patients using urine and sputum samples. *PLoS One.* 2010; **5**(3):e9848.
  10. Blakemore R, Nabeta P, Davidow AL, Vadwai V, Tahirli R, Munsamy V, et al. A multisite assessment of the quantitative capabilities of the Xpert MTB/RIF assay. *American journal of respiratory and critical care medicine.* 2011 Nov 1; **184**(9):1076-84.
  11. Munseri PJ, Talbot EA, Bakari M, Matee M, Teixeira JP, von Reyn CF. The bacteraemia of disseminated tuberculosis among HIV-infected patients with prolonged fever in Tanzania. *Scand J Infect Dis.* 2011 Sep; **43**(9):696-701.
  12. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Clinical significance of lipoarabinomannan (LAM) detection in urine using a low-cost point-of-care diagnostic assay for HIV-associated tuberculosis. *AIDS.* 2012 May 3.
  13. Peter JG, Theron G, Muchinga TE, Govender U, Dheda K. The diagnostic accuracy of urine-based Xpert MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce *PLoS One.* 2012; **7**(7):e39966

References continued from page 42

- The diagnostic yield and safety of fine needle aspiration of intrathoracic Hydatid cysts: A three-year retrospective study**
16. Koegelenberg CF, Bolliger CT, Plekker DP, et al. Diagnostic yield and safety of ultrasound-assisted biopsies in superior vena cava syndrome. *Eur Respir J* 2009; **33**: 1389-1395.
  17. Koegelenberg CFN, Bolliger CT, Irusen EM, et al. The diagnostic yield and safety of ultrasound-assisted transthoracic fine needle aspiration of drowned lung. *Respiration* 2011; **81**:26-31.
  18. Koegelenberg CF, Bolliger CT, Irusen EM, et al. The diagnostic yield and safety of ultrasound-assisted biopsy of mediastinal masses. *Respiration* 2011; **81**: 134-141.
  19. El Kady N, Ramzy I, Hanan HA, Haleem A, El-Bahnasawy MM. Echoguided pair technique in diagnosis and treatment of abdominal hydatid cystic disease in Egyptian patients: clinical and ultrasonographic follow up. *J Egypt Soc Parasitol* 2011; **41**: 527-542.
  20. Eris FN, Akisu C, Aksoy U. Evaluation of two ELISA and two indirect hemagglutination tests for serodiagnosis of pulmonary hydatid disease. *Korean J Parasitol.* 2009; **47**: 427-329.

The South African Respiratory Journal is distributed free of charge.

Should you wish to receive future editions and have not yet completed a registration card, please complete the card in this edition and return it to the address below or email your details to **sarj@iafrica.com**.

The South African Respiratory Journal  
P O Box 16433  
VLAEBERG  
8018

Any correspondence to the Editor should be sent to the same address or via email to:  
**sarj@iafrica.com**

The website of the South African Thoracic Society can be found at:  
**www.pulmonology.co.za**

All articles must be submitted electronically using Editorial Manager. The link to Editorial Manager can be found on the website of the South African Thoracic Society (**www.pulmonology.co.za**).

**Should you require any assistance with article submission please contact the office of the South African Respiratory Journal -  
email: sarj@iafrica.com or telephone: 021-423 0257**

Chapter 9.

**Can point-of-care urine LAM strip testing for tuberculosis add value to clinical decision making in hospitalised HIV-infected persons?**

Peter, J.G., Theron, G. and Dheda, K.

Published in *Plos One* 8(2): e54875

**PhD context**

This final study of the PhD thesis aimed to investigate whether the urine LAM strip test, with modest overall test sensitivity, offered any advantages over routine empiric treatment practices based on a set of basic clinical and radiological predictors. Given our increasing recognition of the attenuating effects of routine empiric treatment practices on the impact of different diagnostic strategies, together with the known high rates of empiric treatment utilised in hospitalised HIV-infected patients with advanced immunosuppression, we hypothesised that urine LAM strip testing may offer little incremental benefit over existing routine empiric treatment practices.

# Can Point-of-Care Urine LAM Strip Testing for Tuberculosis Add Value to Clinical Decision Making in Hospitalised HIV-Infected Persons?

Jonathan G. Peter<sup>1</sup>, Grant Theron<sup>1</sup>, Keertan Dheda<sup>1,2,3\*</sup>

**1** Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, University of Cape Town, Cape Town, South Africa, **2** Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, Cape Town, South Africa, **3** Department of Infection, University College London Medical School, London, United Kingdom

## Abstract

**Background:** The urine lipoarabinomannan (LAM) strip-test (Determine<sup>®</sup>-TB) can rapidly rule-in TB in HIV-infected persons with advanced immunosuppression. However, given high rates of empiric treatment amongst hospitalised patients in high-burden settings (~50%) it is unclear whether LAM can add any value to clinical decision making, or identify a subset of patients with unfavourable outcomes that would otherwise have been missed by empiric treatment.

**Methods:** 281 HIV-infected hospitalised patients with suspected TB received urine LAM strip testing, and were categorised as definite (culture-positive), probable-, or non-TB. Both the proportion and morbidity of TB cases identified by LAM testing, early empiric treatment (initiated prior to test result availability) and a set of clinical predictors were compared across groups.

**Results:** 187/281 patients had either definite- (n = 116) or probable-TB (n = 71). As a rule-in test for definite and probable-TB, LAM identified a similar proportion of TB cases compared to early empiric treatment (85/187 vs. 93/187, p = 0.4), but a greater proportion than classified by a set of clinical predictors alone (19/187; p < 0.001). Thirty-nine of the 187 (21%) LAM-positive patients who had either definite- or probable-TB were missed by early empiric treatment, and of these 25/39 (64%) would also have been missed by smear microscopy. Thus, 25/187 (8%) of definite- or probable-TB patients with otherwise delayed initiation of TB treatment could be detected by the LAM strip test. LAM-positive patients missed by early empiric treatment had a lower median CD4 count (p = 0.008), a higher median illness severity score (p = 0.001) and increased urea levels (p = 0.002) compared to LAM-negative patients given early empiric treatment.

**Conclusions:** LAM strip testing outperformed TB diagnosis based on clinical criteria but in day-to-day practice identified a similar proportion of patients compared to early empiric treatment. However, compared to empiric treatment, LAM identified a different subset of patients with more advanced immunosuppression and greater disease severity.

**Citation:** Peter JG, Theron G, Dheda K (2013) Can Point-of-Care Urine LAM Strip Testing for Tuberculosis Add Value to Clinical Decision Making in Hospitalised HIV-Infected Persons? PLoS ONE 8(2): e54875. doi:10.1371/journal.pone.0054875

**Editor:** Adithya Cattamanchi, San Francisco General Hospital, University of California San Francisco, United States of America

**Received:** October 5, 2012; **Accepted:** December 17, 2012; **Published:** February 4, 2013

**Copyright:** © 2013 Peter et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** JP is supported by a Carnegie Academic Fellowship, a SAMA fellowship and the EDCTP (TB-NEAT). GT is supported by the EDCTP (TB-NEAT), the South African National Research Foundation, and the Claude Leon Foundation. KD is supported by the EDCTP (TB-NEAT, TESA), the SA DST and the SA NRF (SARChI). Alere provided in-kind co-funding in the form of ELISA and LAM strip tests free of charge. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** Keertan Dheda has received speaker fees and in-kind co-funding from Alere in the form of ELISA and LAM strip tests free of charge. Keertan Dheda is a PLOS ONE Editorial Board member. There are no further patents, products in development or marketed products related to this work to declare. This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials, as detailed online in the guide for authors.

\* E-mail: Keertan.Dheda@uct.ac.za

## Introduction

The early high mortality (>25%) amongst hospitalised TB HIV co-infected patients in resource-poor settings requires urgent attention [1,2]. The increased incidence of sputum pauci-bacillary, and disseminated forms of TB in these patients limits the use of both traditional and new TB diagnostic tools [3–6]. Empiric TB treatment, based only on clinical and radiological findings is common (~50%) amongst hospitalised HIV-infected patients with advanced immunosuppression; given their high pre-test probability of disease and illness severity [7,8]. Formalized World Health Organisation (WHO) clinical algorithms are available to guide

empiric treatment and, despite modest diagnostic accuracy in ambulatory patients [9,10], evidence suggests that their use may reduce mortality amongst hospitalised HIV-infected patients [8]. Although death due to undiagnosed TB is common in hospitalised HIV-infected patients in Africa [2,8], empiric treatment guidelines and practices are inconsistent, vary between hospitals, and may needlessly expose patients to toxic drugs. Thus, there remains a need for simple bedside tools to help guide the early initiation of TB treatment, where a microbiological diagnosis may be unavailable or delayed.

Urinary lipoarabinomannan (LAM) has more recently been evaluated for the diagnosis of TB in HIV-infected patients [11,12].

In hospitalised HIV-infected patients, a urinary LAM ELISA (Alere, USA) has an overall sensitivity of 59–67%, increasing to as high as 85% in patients with CD4<50 cells/ml, and an overall specificity of 80–94% [13–15]. In addition, LAM positivity has been associated with higher mycobacterial burden, more severe illness, and a higher mortality [15–17]. As an alternative to the ELISA-kit, LAM can now be detected by a simple, low-cost (<US\$3.5) point-of-care lateral flow assay that is able to provide results in 25 min from just 60 µl of urine [11,12]. Initial evaluation studies have confirmed equivalent performance of the urine LAM strip test compared to the LAM ELISA assay in different settings [13,18].

However, is this test really useful in ‘real world’ clinical practice in high HIV prevalence settings? The real value in any diagnostic lies in its ability to provide information beyond that deducible from basic clinical and radiographic data, such that it adds incremental value to routine clinical practice. Useful tests add value to clinical decision-making by ruling-in patients not otherwise routinely identifiable, pinpointing otherwise unrecognizable patients with the highest risk of morbidity and mortality, or by reducing unnecessary treatment. In this context, our study investigated whether point-of-care urine LAM strip testing offered any value over basic clinical and radiological screening, and whether testing was redundant in the context of routine ‘real world’ day-to-day clinical practice where empiric treatment is commonly used. We therefore evaluated LAM strip test performance against physician-led empiric treatment decisions and a set of clinical predictors.

## Methods

### Study Population

A study outline is shown in Figure 1. In total, 335 prospectively recruited adult in-patients patients from four hospitals (three district- and one tertiary-level) between July 2009 and December 2010 in Cape Town, South Africa were enrolled. Patients were referred for study inclusion by emergency-room or clinic doctors if suspected to have HIV-TB co-infection and needed in-patient care. All patients provided written informed consent and the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee approved the study. Clinical information collected included: demographics, past history of TB, co-morbidity, symptoms, vital signs (including weight) and a modified early warning (MEWS) illness severity score [19]. Blood was taken for HIV, CD4 and renal function testing. A chest radiograph (CXR) was performed in all patients.

### TB Diagnostic Sampling and Testing

Consultant-led hospital-based clinicians not associated with the study team determined the timing and extent of TB diagnostic work-up and the commencement of empiric anti-TB treatment. Routine hospital practice includes the collection, where possible, of two sputum samples in patients able to expectorate and, if extra-pulmonary TB is suspected, the collection of 1–2 non-sputum samples from clinically involved sites (e.g. fine needle aspirate of lymph node, pleural fluid aspirate/biopsy, ascitic tap, lumbar puncture, pericardial aspiration etc.). Further details of biological samples collected for TB culture, stratified by patient TB diagnostic category, are outlined in table 1. Concentrated fluorescence smear microscopy was performed on NALC/NaOH processed sputum/non-sputum samples, and cultures were performed using the MGIT 960 liquid culture system (BD Diagnostics, USA).

### TB Reference Standard and Case Definitions

The reference standard for definite-TB was liquid culture positivity for *Mycobacterium tuberculosis* from at least a single sample. Given the significant potential for misclassification bias due to challenges of sampling extra-pulmonary compartments, the significant proportion of sputum scarce patients, and the limited performance of a single liquid TB culture in HIV-infected patients [20], patients were further categorised into the following diagnostic groups for analysis (Figure 1):

**Definite-TB.** At least 1 *M. tuberculosis* sample positive by liquid culture (either sputum or non-sputum e.g. pleural fluid, pericardial fluid etc.).

**Probable-TB.** Not meeting the criterion for definite-TB but a clinical-radiological picture highly suggestive of TB. All patients in this group received and showed a good response to anti-TB treatment at two-month follow-up. Smear-positive but culture-negative or contaminated patient samples were included in this group.

**Non-TB.** No culture-based evidence of *M. tuberculosis* and an alternative diagnosis available. No clinical deterioration on two-month follow-up and no TB treatment given. Patients culture positive for non-tuberculosis mycobacteria (NTM) and not receiving anti-TB treatment were assigned to this group.

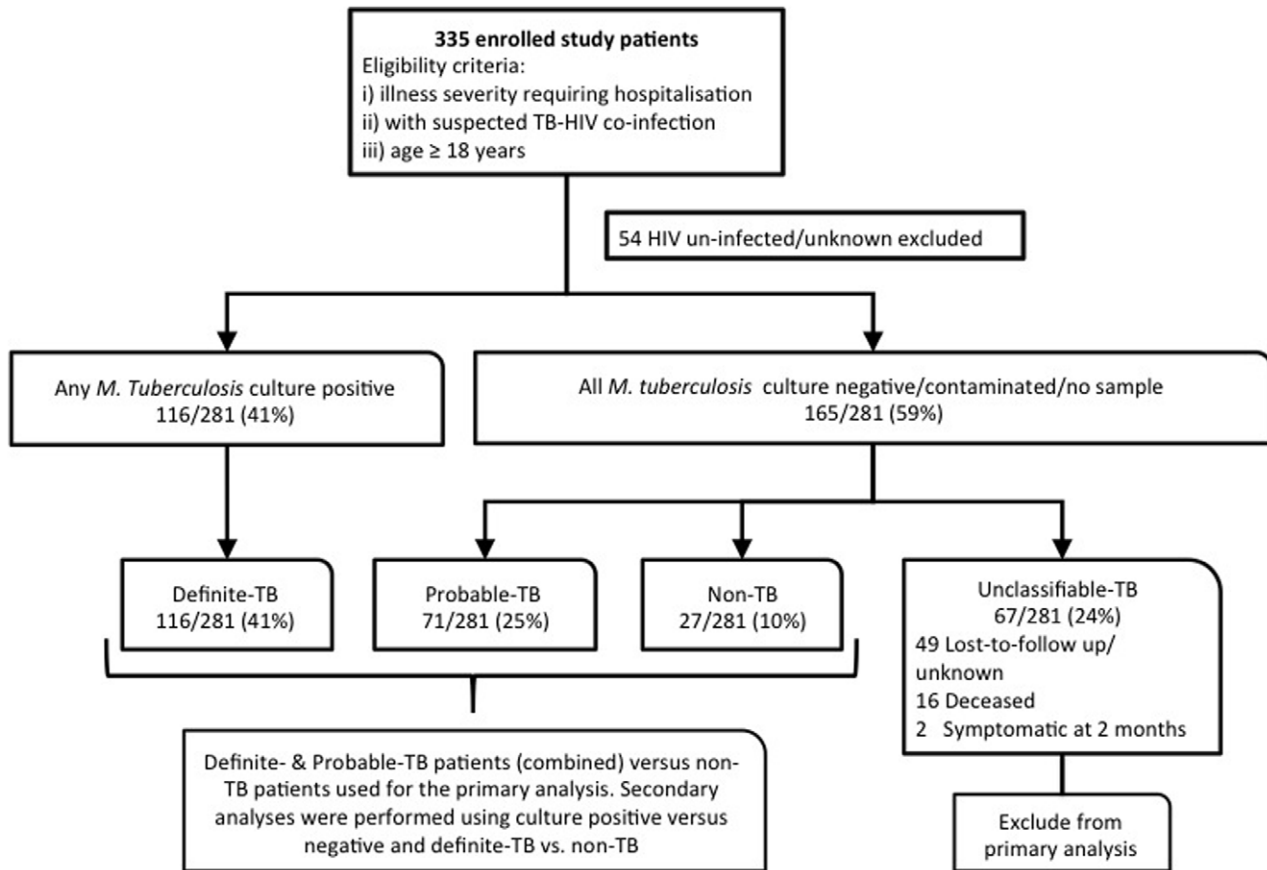
**Unclassifiable TB.** Unable to assign to any of the above-mentioned diagnostic groups due to death of unknown cause (without autopsy), on-going but uncharacterised symptoms at follow-up, or loss-to-follow-up at 2 months.

### Early Empiric Treatment Definition

In order to compare the diagnostic performance of the urine LAM strip testing with routine clinical practice, **early empiric treatment** was defined as any patient commencing TB treatment within 24 hours of hospital admission based only on clinical and/or radiological findings, and prior to the availability of any smear or culture results. All early empiric treatment decisions, even if initial made by junior staff (medical officers and registrars), were approved by the attending consultant general physician.

### Modelling Clinical Predictors Using Multiple Imputation

A univariate analysis was used to determined basic clinical, laboratory and radiological predictors of definite-TB. A set of multivariate clinical predictors was generated using stepwise logistic regression modeling. Multiple imputation by chained Equations (Royston, P & White, I 2011) was used to impute missing data prior to model building. The variables included in the logistic regression modeling included the following (number of missing data points for each variable that were imputed is indicated in the brackets): sex (2), age (5), previous TB history (0), known TB contact (0), current smoker (0), cough  $\geq 2$  weeks (0), productive cough (0), haemoptysis (0), self-reported weight loss (0), appetite loss (0), recent fever (0), night sweats (0), fatigue (0), shortness of breath (0), chest pain (0), abdominal pain (0), nausea/vomiting (0), diarrhea (0), neurological symptoms (0), measured weight (39), temperature (7), respiratory rate (7), and Modified early warning score (MEWS) (144) at enrollment, CXR compatibility with TB (24), and urine dipstick abnormalities (0). Different data appeared to be missing for different patients in a random fashion. The continuous variables (weight and temperature) were dichotomised using receiver operating characteristic (ROC) analysis to identify cut-points that maximised discriminatory utility prior to inclusion in the model. Rounded  $\beta$ -coefficients from the reduced model of significant variables were used to generate scores to quantitate relevant clinical predictors. ROC analysis was performed and three cut-points were selected for rule-



**Figure 1. Study population, culture status and diagnostic groups used for analyses.**  
doi:10.1371/journal.pone.0054875.g001

in, Youden's index (the optimal mathematical balance between sensitivity and specificity [21]) and rule-out value. Diagnostic accuracy, including 95% confidence intervals, for each cut-point was assessed using sensitivity, specificity, positive and negative predictive values (PPV, NPV) and positive likelihood ratio (LR+).

### Urine Sampling and LAM Methodology

All patients gave a spot urine sample (10–30 ml) collected in a sterile container as soon as possible after recruitment. Urine was frozen on the day of collection and stored at  $-20^{\circ}\text{C}$  for later batched testing. Urine LAM strip testing (Determine<sup>®</sup> TB, Alere, USA) was performed on thawed urine according to the manufacturer's instructions by readers blinded to all patient data and reference test results. Urine LAM strip test lot#101102, the same as used for test evaluation in an outpatient ARV-clinic setting [18], was used. Detailed methodology for reading the urine LAM strip tests has been previously described [13]. Analysis was performed using the grade 2 cut-point which has shown better inter-observer reliability and good rule-in value ( $\text{LR+} > 10$ ) in hospitalised HIV-infected patients [13].

### Statistical Analysis

Analyses were restricted to HIV infected patients only and were performed using definite- and probable- (combined) versus non-TB patient groups for the primary determination of diagnostic accuracy (unclassifiable patients were excluded). In addition, given the inability to accurately evaluate the specificity of empiric treatment in the primary analysis as treatment response formed

part of the diagnostic categorisation, alternative analyses were performed (see online supplementary material) using either *M. tuberculosis* complex culture-positive versus negative groups, or only definite- versus non-TB groups. Diagnostic accuracy, including 95% confidence intervals, for individual tests and early empiric treatment was assessed using sensitivity, specificity and LR+. Given the variations in test specificity and very high study prevalence of TB, ranges of positive and negative predictive values are presented for individual tests and early empiric treatment at differing rates of in-patient TB prevalence. STATA IC, version 11 (Stata Corp, Texas, USA) was used for all statistical analyses.

## Results

### Demographics, Basic Clinical and TB Diagnostic Sampling Characteristics

Figure 1 outlines the study population. 16% (54/335) of enrolled patients were HIV uninfected and hence excluded from further analysis. 41% (116/281) of patients had definite-TB, an additional 25% (71/281) of patients had probable-TB, and only 10% (27/281) had non-TB. 24% (67/281) of patients remained unclassifiable due to death or lost-to-follow-up and were excluded from the primary analysis. Table 1 outlines demographics, basic clinical characteristics of the patient cohort and the sputum/non-sputum diagnostic samples stratified by TB diagnostic category. These same patient characteristics stratified by smear, culture and CD4 count have been previously described [13,22]. The median

**Table 1.** Demographic, clinical, sampling and microbiological characteristics of study patients stratified by TB diagnostic group.

	All N = 281	Definite TB N = 116	Probable TB N = 71	Non TB N = 27	Unclassified TB N = 67	P-value
<b>Demographics</b>						
Median age (yrs, IQR)	35 (29–39)	35 (28–39)	33 (30–39)	34 (32–40)	36 (30–41)	n/s
Female	174 (62)	69 (59)	45 (63)	19 (70)	41 (61)	n/s
Median CD4 count (cells/ml, IQR)	89 (46–198)	86 (42–192)*	120 (51–215)*	128 (74–235)*	77 (42–120)*	*0.001
Previous TB	97 (35)	33 (28)*	22 (31)	11 (41)	31 (46)*	*0.01
Current Smoker	52 (19)	25 (22)	9 (13)	3 (11)	15 (22)	n/s
<b>Clinical features</b>						
Cough >2wks	231 (82)	99 (85)	58 (82)	22 (82)	52 (78)	n/s
Night sweats	185 (66)	82 (71)	49 (69)	14 (52)	40 (60)	n/s
Weight loss	247 (88)	105 (91)	61 (86)	22 (82)	59 (88)	n/s
Fever >38°C	49 (18)	29 (26)* <sup>†1</sup>	9 (13)*	3 (12)	8 (12) <sup>†1</sup>	*0.04 <sup>†1</sup> 0.03
CXR compatible with TB	215 (77)	99 (85)* <sup>†1</sup>	50 (70)*	16 (60) <sup>†1</sup>	50 (75)	*0.01 <sup>†1</sup> 0.002
Early Empiric Rx given	120 (43)	59 (51)	34 (48)	n/a	27 (40)	n/s
LAM strip test positive (grade 2)	98 (35)	58 (50)*	27 (38) <sup>†1</sup>	1 (4)	12 (18)* <sup>†1</sup>	*<0.001 <sup>†1</sup> 0.009
<b>Clinical samples collected for TB culture</b>						
1 sputum sample	207 (74)	95 (82)*	43 (61)*	21 (78)	48 (72)	*0.001
≥2 sputum samples	92 (33)	42 (26)	16 (23)	9 (33)	25 (37)	n/s
1 non-sputum sample	160 (57)	78 (67)* <sup>†1</sup>	43 (61)	11 (41)*	28 (42) <sup>†1</sup>	*0.01 <sup>†1</sup> <0.001
≥2 non-sputum sample	56 (20)	26 (22)	14 (20)	3 (11)	13 (19)	n/s
No samples	19 (7)	0 (0)	5 (7)	4 (15)	10 (15)	n/s

P-values indicate significant differences between patient groups (marked with \* and number to indicate comparison group) for different patient characteristics. n/a: not applicable; n/s: not significant (p>0.05).

doi:10.1371/journal.pone.0054875.t001

(IQR) CD4 cell count of unclassifiable patients was 77 (42–120) cells/ml, similar to definite-TB patients [86 (42–192) cells/ml], but lower than either probable- [120 (51–215) cells/ml] or non-TB patients [128 (74–235) cells/ml] (p = 0.001). Compared to definite-TB patients, unclassifiable patients provided fewer non-sputum samples for TB culture [67% (78/116) vs. 42% (28/67), p<0.001].

### Routine Early Empiric Treatment Compared to the Urine LAM Strip Test

Table 2 compares the sensitivity (95% CI), specificity and LR+ of routine early empiric treatment, urine LAM strip test and CXR for definite- and probable-TB compared to non-TB patients (definite-TB to non-TB patients only compared in table S2). Early empiric treatment identified and the urine LAM strip test diagnosed an approximately equal number TB cases with sensitivities of 50% (43–57, 93/187) and 46% (39–53, 85/187) respectively (p = 0.4). The combined sensitivity of urine LAM strip testing together with early empiric treatment was higher than either modality alone [LAM and early empiric treatment: 71% (95% CI: 64–77, 132/187) vs. LAM alone: 46% (95% CI: 39–53, 85/187) vs. early empiric treatment alone: 50% (95% CI: 43–57, 93/187), both p<0.001]. In the primary analysis, both early empiric treatment and the urine LAM strip test showed specificities >95%. However, when an alternative analysis restricted to only patients with a valid *M. tuberculosis* culture result and comparing culture-positive and -negative groups is performed

(results provided in table S1), although test sensitivities are not significantly lower, the specificities of both early empiric treatment and urine LAM strip testing decrease, with urine LAM offering higher specificity than early empiric treatment [75% (67–82, 95/126) vs. 63% (54–71, 79/126), p = 0.03]. Given this variable test specificity and high overall study TB prevalence, Table 3 presents a range of PPV (95% CI) and NPV values using three specificities for each diagnostic method (lowest, highest and average) at in-patient TB prevalence rates of 35%, 45%, and 55%, which could be expected to occur in the majority of endemic country hospital settings. The lowest specificities used in Table 3 are taken from the specificities presented in table S1, the highest from Table 2 and the third is an average of the highest and lowest. With an in-patient TB prevalence of 45% (*M. tuberculosis* culture positive TB prevalence in study = 48%, 116/242), the PPV ranges for early empiric treatment, urine LAM strip test and a combination thereof was 53–100%, 62–90% and 71–94%, respectively.

### Clinical Predictors Compared to the Urine LAM Strip Test

The univariate and multivariate associates of definite-TB are shown in Table S3. Table S4 shows the sensitivity (95% confidence intervals), specificity, and LR+ for ROC-selected cut-points, selected for their rule-in, rule-out, or best compromise between sensitivity and specificity (assuming equal weighting) for the quantified set of clinical predictors, the urine LAM strip test and early empiric treatment. At equivalent specificity, clinical

**Table 2.** Sensitivity, specificity and positive likelihood ratio of early empiric treatment, the urine LAM strip test, and CXR for TB diagnosis in hospitalised HIV-infected patients using the definite and probable-TB groups for sensitivity, and the non-TB groups for specificity analyses.

Diagnostic method	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	LR+ (95% CI)
Early empiric Rx <sup>†</sup>	50 <sup>*1*2</sup> (43–57) 93/187	100 (88–100) 27/27	N/C
Urine LAM (grade 2 cut-point)	46 <sup>*3*4</sup> (39–53) 85/187	96 <sup>*5</sup> (82–99) 26/27	12.3 (1.7–89.6)
CXR	85 <sup>*1*3</sup> (79–89) 159/187	30 <sup>*5</sup> (16–49) 8/27	1.2 (1.1–1.3)
Early empiric Rx plus urine LAM (grade 2 cut-point)	71 <sup>*2*4</sup> (64–77) 132/187	96 (82–99) 26/27	19.1 (2.7–136.1)

<sup>†</sup>Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group.

P-values indicate significant differences between tests (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures.

\*1p<0.001;

\*2p<0.001;

\*3p<0.001;

\*4p<0.001;

\*5p<0.001;

\*6p = 0.006.

LAM: Lipoarabinomannan; CXR: Chest X-ray; LR+: positive likelihood ratio; 95% CI: 95% Confidence interval; Rx: treatment.

doi:10.1371/journal.pone.0054875.t002

predictors  $\geq 2.5$  had a lower sensitivity than urine LAM strip testing [10% (95% CI: 7–15, 19/187) vs. 45% (95% CI: 38–53, 85/187),  $p < 0.001$ ].

### Clinical Predictors and Early Empiric Treatment

42% (10/24) of patients ‘ruled-in’ by clinical predictors  $\geq 2.5$  and 38% (23/61) patients ‘ruled-out’ by clinical predictors  $\leq 0.5$  were given early empiric treatment by attending hospital clinicians. Table S5 provides a further comparison of patient characteristics for patients commencing vs. not commencing early empiric treatment. No differences in basic demographic, symptomatology or diagnostic sampling was noted between groups except that a higher proportion of patients given early empiric treatment had a cough  $> 2$  weeks [90% (108/120) vs. 76% (123/161),  $p = 0.003$ ]. Patients given vs. not given early empiric treatment had a lower median MEWS [3 (1–5) vs. 4 (3–5),  $p = 0.001$ ] and creatinine level [68 (56–94) vs. 77 (59–107)  $\mu\text{mol/l}$ ,  $p = 0.04$ ].

### Urine LAM Strip Test Positive Patients Missed by Early Empiric Treatment

The Venn diagram in Figure 2 indicates the different but overlapping patient populations detected by urine LAM strip and early empiric treatment initiation. 21% (39/187) of definite- and probable-TB cases were urine LAM strip test positive, but missed by early empiric treatment. 64% (25/39) of these patients were either sputum smear-negative or unable to produce sputum. Table 4 compares the characteristics of definite- and probable-TB patients detected by either urine LAM alone, early empiric treatment or missed by both modalities. Patients detected by urine LAM alone, when compared to those detected by early empiric treatment alone, had a lower median (IQR) CD4 cell count [73 (31–134) vs. 138 (60–217),  $p = 0.008$ ], a higher median (IQR) MEWS [5 (3–6) vs. 3 (1–4),  $p = 0.001$ ] and an increased urea level [5.6 (4.0–11.6) vs. 3.7 (3.2–4.6),  $p = 0.002$ ]. Patients detected by

urine LAM alone, when compared to those missed by both urine LAM and early empiric treatment, had lower median CD4 cell count [73 (31–134) vs. 173 (58–269) units,  $p = 0.003$ ]. If the analysis was repeated using only definite-TB patients the findings were similar (Figure S1) and conclusions unchanged. 24% (28/116) of culture-positive patients were LAM positive but missed by early empiric treatment. When compared to the 23% (27/116) of patients that were LAM-negative but started on early empiric treatment, they had a lower median (IQR) CD4 cell count [62 (31–110) vs. 107 (54–171) cells/ml,  $p = 0.05$ ], a higher median (IQR) MEWS [5 (3–6) vs. 3 (1–4),  $p = 0.01$ ] and an increased urea level [5.5 (3.7–11.2) vs. 3.7 (3.4–4.6),  $p = 0.02$ ]. Patients detected by urine LAM alone had lower median (IQR) CD4 cell count and a higher 8-week mortality than those patients missed by both urine LAM and early empiric treatment [CD4 count: 62 (31–110) vs. 180 (74–308),  $p = 0.002$ ; 8-week mortality: 24% (6/25) vs. 4% (1/23),  $p = 0.05$ ].

### Discussion

The point-of-care urine LAM test has potential as a useful adjunct for rapid TB diagnosis in HIV-infected hospitalised patients [11,13]. Its added clinical value, however, remains uncertain given its modest performance characteristics. The key finding of this study is that LAM detected patients that would have otherwise been missed by empiric treatment and this subgroup of patients had more advanced immunosuppression and greater illness severity. The latter represents a group most likely to benefit from the initiation of early treatment as they are at high risk.

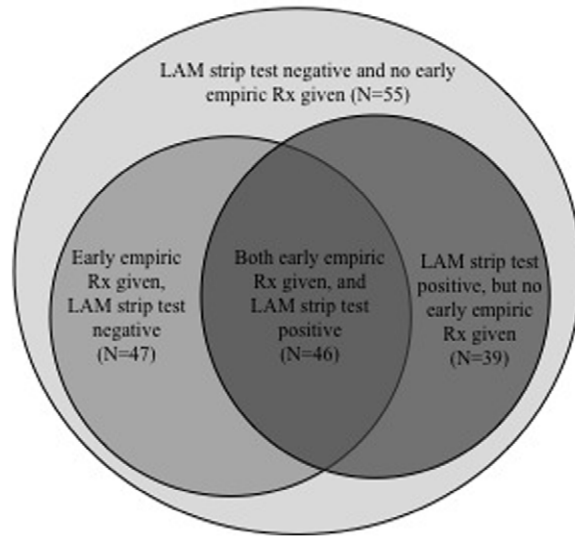
Traditional and newer TB diagnostics show reduced diagnostic accuracy in hospitalised co-infected patients, particularly with advanced immunosuppression, as patients are often unable to produce sputum for diagnostic testing and/or have disseminated disease [3]. In addition, these patients present to hospitals with late stage disease and severe illness [3]. These factors mean that treatment decisions are commonly made based on clinical and

**Table 3.** Calculated positive and negative predictive values for early empiric treatment, the urine LAM strip test and a combination thereof, using three data-generated estimates for test specificity and in-patient TB prevalence.

Diagnostic	Test sensitivity (%)	Test specificity (%)	In-patient TB prevalence		
			35%	45%	55%
Early Empiric Rx	51	63	43 (38-47)	53 (48-58)	63 (58-67)
	51	82	60 (55-66)	70 (65-75)	78 (73-82)
	50	100	100 (98-100)	100 (98-100)	100 (99-100)
Urine LAM strip test	50	75	52 (47-57)	62 (57-70)	71 (66-75)
	48	86	65 (59-70)	74 (68-78)	81 (76-85)
	46	96	86 (80-90)	90 (86-94)	93 (90-96)
Early empiric Rx plus urine LAM (grade 2 cut-point)	74	75	61 (57-66)	71 (67-75)	78 (75-82)
	73	86	74 (69-78)	81 (77-85)	87 (83-89)
	71	96	91 (87-94)	94 (90-96)	96 (93-97)

LAM: Lipoarabinomannan; PPV: positive predictive value; NPV: negative predictive value; 95% CI: 95% Confidence interval; Rx: treatment.

All definite and probable TB patients (N=187)



**Figure 2.** Venn diagram of all definite- and probable-TB patients indicating different but overlapping patient populations detected by the urine LAM strip test and early empiric TB treatment.

doi:10.1371/journal.pone.0054875.g002

radiological findings alone, the need for urgent treatment initiation, and the high background disease prevalence (pre-test probability). Yet, in this same patient group, clinical and radiological findings are frequently atypical and/or non-specific, and this accounts for the poor rule-in value of the set of clinical predictors that we derived. Indeed, as a ‘rule-in’ test (cut-point selecting high specificity and PPV) clinical predictors could only correctly classify ~20% of all patients. This assumes coherent mathematical analysis of available diagnostic variables. However, routine clinical decision-making is rather a dynamic Bayesian process of assimilating an accumulating series of pre-test probabilities [23], and weighting of the overall post-test probability of disease against a threshold probability for initiating treatment. Thus, in reality physician practice varies widely and in an attempt to reduce mortality hospital treatment thresholds are lower than expected. Indeed, in this study, ~50% of definite- and probable-TB patients initiated early empiric TB treatment.

Early empiric TB treatment decisions should logically be targeted to the sickest patients, especially amongst hospitalised HIV-infected patients with advanced immunosuppression where given higher mortality rates, treatment-initiation thresholds should be lowered. However, this was not the case in those empirically treated. In fact overall, patients not commenced on early empiric treatment appeared, using the MEWS, to have a higher illness severity. No clear demographic, clinical or radiological factor predicted early empiric treatment. By contrast, the urine LAM strip test could pinpoint the most severely ill patients. Thus, our data suggest that early empiric treatment will miss a particularly vulnerable patient group with advanced immunosuppression that would have been detected by LAM strip testing. The rapid identification of these patients could target prompt therapy to those most likely to benefit. Recent studies with both the TB LAM ELISA and the LAM strip test support our findings, demonstrating associations between urine LAM positivity, higher mycobacterial disease burden, more advanced immunosuppression, and increased illness severity and mortality [15–17]. These findings

**Table 4.** Comparison of patient characteristics between groups detected by either the urine LAM strip or early empiric treatment alone, or missed both bedside diagnostic methods (only patients with definite- or probable-TB were included).

Patient characteristic(s)	LAM strip positive, but no early empiric Rx given	LAM strip negative, but early empiric Rx given	LAM strip negative and no early empiric treatment given	†P-value
	(N = 39)	(N = 47)	(N = 55)	
	Median (IQR) or n(%)	Median (IQR) or n(%)	Median (IQR) or n(%)	
Age, years	32 (26–38)	35 (28–38)	35 (30–40)	n/s
Previous TB	12 (31)	16 (34)	18 (33)	n/s
Weight, kg	52 (47–60)	56 (44–63)	52 (47–61)	n/s
Temperature, °C	36.9 (36.0–37.1)	36.9 (36.5–38.0)	36.9 (36.3–37.7)	n/s
Respiratory rate, breaths/min	24 (20–28)	22 (19–29)	24 (18–28)	n/s
CD4 cell count, cells/ml	73 (31–134)* #	138 (60–217)*	173 (58–269)#	*0.008 #0.003
MEWS <sup>‡</sup>	5 (3–6)*	3 (1–4)*	4 (3–6)	*0.001
Urea, mmol/l	5.6 (4.0–11.6)*	3.7 (3.2–4.6)*	4.2 (3.2–7.8)	*0.002
Creatinine, µmol/l	95 (59–121)*	67 (55–79)*	74 (57–102)	*0.01
8-week mortality	6/36 (17)	3/40 (8)	3/39 (8)	n/s

<sup>‡</sup>MEWS: Modified early warning score is an admission triage score based on illness severity and higher scores correlated with poor outcomes and increased mortality [19].

†P-values indicate significant differences between patient groups (marked with \* and # to indicate comparison group) for different patient characteristics.

LAM: Lipoarabinomannan; Rx: treatment; IQR: interquartile range.

doi:10.1371/journal.pone.0054875.t004

support the need to undertake prospective impact studies to assess whether initiation of TB treatment based on urine LAM testing is able to save lives and/or decrease TB-related morbidity.

Do our findings have relevance to in-patient settings with a lower TB prevalence? Amongst in-patient settings with a lower TB prevalence, urine LAM is likely to offer superior 'rule-in' utility compared to empiric treatment. This is evidenced by: i) the poor comparative diagnostic utility and inferiority of a set of clinical predictors, which estimates pre-test probability or what could be expected with a frequentist interpretation of simple clinical and radiological predictors, and ii) lack of clear demographic, clinical or radiological parameters associated with early empiric treatment practice and limited agreement with the derived set of clinical predictors indicating the lack of predictability and hence, standardisation of empiric treatment decision-making. Given the modest performance characteristics of the urine LAM strip testing it is however clear that both urine LAM alone, or combined with existing empiric treatment practise, is likely to only offer clinically useful 'rule-in' utility (PPV > 90%) in hospital settings with high TB prevalence (> 35%).

This is the first study to compare the value of urine LAM strip testing against clinical-radiological screening and day-to-day clinical practice (early empiric treatment rates) in hospitalised patients. However, our study has important limitations. Given the well-established misclassification bias that occurs due to the drawbacks of the TB culture technique and the lack of interventional lung sampling (sputum induction and bronchoscopy), a diagnostic categorisation was used to group patients for analysis. This may have underestimated sensitivity and overestimated specificity. The TB prevalence in our study was higher than in many other settings and this limits the generalisability of our findings. However, we have presented predictive values using estimated low, medium and high TB prevalence rates to improve generalisability. The study had a high proportion of unclassifiable-TB patients due to death and loss-to-follow-up, and these patients had a higher proportion of LAM strip test negative results than

definite- and probable-TB groups. This may have introduced the possibility of selection bias, however, in our secondary analyses presented in the online supplementary materials we compare *M. tuberculosis* culture positive vs. negative groups and include unclassifiable-TB patients with a valid culture result. Key study findings are unaffected. Our study did not evaluate LAM against newer diagnostic standards such as the Xpert MTB/RIF assay. However, this was not accessible to us at the time of the study and this test offers reduced utility in extra-pulmonary and sputum scare TB. Urine LAM test results were not performed at the bedside or used to guide treatment initiation, thus a survival benefit through initiating early treatment in these severely ill patients is unclear but possible [8].

An ideal point-of-care test for rapid, laboratory-free detection of TB remains elusive [24]. However sampling hurdles and poor performance when using extra-pulmonary samples mean that the need to make empiric treatment decisions is likely to continue. Thus, despite only modest diagnostic accuracy, the low cost urine LAM strip test offers important added clinical value in hospitalised HIV-infected patients with suspected TB. Not only could the test detect patients missed by clinical and radiological predictors but also could potentially enable the rapid treatment of patients with the most advanced immunosuppression and severe illness. Further studies are now required to confirm our study findings and evaluate the impact of urine LAM strip testing to guide early treatment initiation in hospitalised HIV-infected patients.

### Supporting Information

**Figure S1** Venn diagram of all definite-TB patients indicating different but overlapping patient populations detected by the urine LAM strip test and early empiric TB treatment. (TIFF)

**Table S1** Diagnostic accuracy measures of early empiric treatment, the urine LAM strip test and CXR for TB diagnosis in hospitalised HIV-infected patients using *M. tuberculosis* culture

positive-TB patients for sensitivity, and culture negative patients for specificity analyses. <sup>‡</sup>All patients with 1 or more valid *M. tuberculosis* culture (either sputum or non-sputum) are included in this analysis irrespective of final TB diagnostic categorization (39/281 patients excluded with either no/contaminated culture result) <sup>†</sup>Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group. P-values indicate significant differences between tests (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures <sup>\*1</sup>p<0.001; <sup>\*2</sup>p<0.001; <sup>\*3</sup>p<0.001; <sup>\*4</sup>p<0.001; <sup>\*5</sup>p = 0.03; <sup>\*6</sup>p = 0.03; <sup>\*7</sup>p<0.001; <sup>\*8</sup>p = 0.005. (DOCX)

**Table S2** Diagnostic accuracy measures of early empiric treatment, the urine LAM strip test and CXR for TB diagnosis in hospitalized HIV-infected patients using definite-TB (*M. tuberculosis* culture positive) for sensitivity, and non-TB patient groups for specificity analyses. <sup>†</sup>Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group. P-values indicate significant differences between tests (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures <sup>\*1</sup>p<0.001; <sup>\*2</sup>p<0.001; <sup>\*3</sup>p<0.001; <sup>\*4</sup>p<0.001; <sup>\*5</sup>p<0.001. (DOCX)

**Table S3** Univariate and multivariate analyses for associates of definite-TB in HIV-infected hospitalised patients. <sup>†</sup>Receiver operating characteristic (ROC) curve-selected cut-point maximizing discriminatory utility used to dichotomise the continuous variables weight and temperature OR: odds ratio; TB: Tuberculosis; CXR: Chest x-ray; LAM: Lipoarabinomannan. (DOCX)

**Table S4** Diagnostic accuracy measures for set of clinical predictors (using three ROC-selected cut-points), the urine LAM strip test and routine early empiric treatment in hospitalised HIV-

infected patients using the definite and probable-TB groups for sensitivity and the non-TB groups for specificity analyses. P-values indicate significant differences between tests and/or cut-points (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures; <sup>\*1</sup>p<0.001; <sup>\*2</sup>p = 0.03<sup>†</sup> Youden's index is defined as the point on the ROC curve that provides the optimal mathematical balance between sensitivity and specificity. (DOCX)

**Table S5** Demographic, clinical, sampling and microbiological characteristics of study patients stratified by TB diagnostic group <sup>†</sup> Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group. Analysis is performed for all patients in this graph and hence includes 27 unclassified patients whom were commenced on early empiric treatment but do not form part of the primary analysis presented in the main manuscript. P-values indicate significant differences between patient groups (marked with \* and number to indicate comparison group) for different patient characteristics <sup>‡</sup>MEWS: Modified early warning score is an admission triage score based on illness severity and higher scores correlated with poor outcomes and increased mortality [19]. (DOCX)

## Acknowledgments

We thank the research nursing staff, including Sr Nicola Kelly and Faranaaz Embrahim, who were involved in the recruitment of patients and collection of urine samples. The support of the clinical and nursing staff of the general medical wards of Groote Schuur, Victoria, GF Jooste and New Somerset hospitals was greatly appreciated.

## Author Contributions

Conceived and designed the experiments: JP KD. Analyzed the data: JP. Contributed reagents/materials/analysis tools: GT. Wrote the paper: JP KD GT.

## References

- Edginton ME, Wong ML, Phofa R, Mahlaba D, Hodkinson HJ (2005) Tuberculosis at Chris Hanani Baragwanath Hospital: numbers of patients diagnosed and outcomes of referrals to district clinics. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease 9: 398–402.
- Cohen T, Murray M, Wallengren K, Alvarez GG, Samuel EY, et al. (2010) The prevalence and drug sensitivity of tuberculosis among patients dying in hospital in KwaZulu-Natal, South Africa: a postmortem study. PLoS Med 7: e1000296.
- Kwan CK, Ernst JD. (2011) HIV and tuberculosis: a deadly human syndemic. Clin Microbiol Rev 24: 351–376.
- Theron G, Peter J, Dheda K (2011) Xpert MTB/RIF test for tuberculosis. Lancet 378: 481; author reply 482–483.
- Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, et al. (2011) Evaluation of the Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting. American journal of respiratory and critical care medicine 184: 132–140.
- Peter J, van Zyl Smit R, Deninger C, Pai M (2012) Diagnosis of tuberculosis: state-of-the-art. European Monograph Series: Tuberculosis. Editors Lange C and Migliori GB European Respiratory Society 58: 124–143
- Bakari M, Arbeit RD, Mtei L, Lyimo J, Waddell R, et al. (2008) Basis for treatment of tuberculosis among HIV-infected patients in Tanzania: the role of chest x-ray and sputum culture. BMC Infect Dis 8: 32.
- Holtz TH, Kabera G, Mthiyane T, Zingoni T, Nadesan S, et al. (2011) Use of a WHO-recommended algorithm to reduce mortality in seriously ill patients with HIV infection and smear-negative pulmonary tuberculosis in South Africa: an observational cohort study The Lancet infectious diseases 11: 533–540.
- Wilson D, Mbhele L, Badri M, Morroni C, Nachege J, et al. (2011) Evaluation of the World Health Organization algorithm for the diagnosis of HIV-associated sputum smear-negative tuberculosis. Int J Tuberc Lung Dis 15: 919–924.
- Koole O, Thai S, Khun KE, Pe R, van Griensven J, et al. (2011) Evaluation of the 2007 WHO guideline to improve the diagnosis of tuberculosis in ambulatory HIV-positive adults. PLoS One 6: e18502.
- Peter J, Green C, Hoelscher M, Mwaba P, Zumla A, et al. (2010) Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments. Curr Opin Pulm Med 16: 262–270.
- Lawn SD. (2012) Point-of-care detection of lipoarabinomannan (LAM) in urine for diagnosis of HIV-associated tuberculosis: a state of the art review. BMC infectious diseases 12: 103.
- Peter JG, Theron G, van Zyl-Smit R, Haripersad A, Mottay L, et al. (2012) Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients. The European respiratory journal : official journal of the European Society for Clinical Respiratory Physiology 40(5): 1211–20
- Shah M, Variava E, Holmes CB, Coppin A, Golub JE, et al. (2009) Diagnostic accuracy of a urine lipoarabinomannan test for tuberculosis in hospitalised patients in a High HIV prevalence setting. Journal of acquired immune deficiency syndromes 52: 145–151.
- Talbot E, Munseri P, Teixeira P, Matee M, Bakari M, et al. (2012) Test characteristics of urinary lipoarabinomannan and predictors of mortality among hospitalized HIV-infected tuberculosis suspects in Tanzania. PLoS One 7: e32876.
- Shah M, Martinson NA, Chaisson RE, Martin DJ, Variava E, et al. (2010) Quantitative analysis of a urine-based assay for detection of lipoarabinomannan in patients with tuberculosis. J Clin Microbiol 48: 2972–2974.
- Lawn SD, Kerkhoff AD, Vogt M, Wood R (2012) Clinical significance of lipoarabinomannan (LAM) detection in urine using a low-cost point-of-care diagnostic assay for HIV-associated tuberculosis. AIDS 26(13): 1635–43
- Lawn SD, Kerkhoff AD, Vogt M, Wood R (2012) 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 infectious diseases* 12: 201–209.
19. Subbe CP, Kruger M, Rutherford P, Gemmel L (2001) Validation of a modified Early Warning Score in medical admissions. *Qjm*,94: 521–526.
  20. Monkongdee P, McCarthy KD, Cain KP, Tasaneeyapan T, Nguyen HD, et al. (2009) Yield of acid-fast smear and mycobacterial culture for tuberculosis diagnosis in people with human immunodeficiency virus. *Am J Respir Crit Care Med* 180: 903–908.
  21. Schisterman EF, Perkins NJ, Liu A, Bondell H (2005) Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples. *Epidemiology* 16: 73–81.
  22. Peter JG, Theron G, Muchinga TE, Govender U, Dheda K (2012) The diagnostic accuracy of urine-based Xpert MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce *PLoS ONE* 7: e39966.
  23. Gill CJ, Sabin L, Schmid CH (2005) Why clinicians are natural bayesians. *BMJ* 330: 1080–1083.
  24. Pai NP, Pai M (2012) Point-of-care diagnostics for HIV and tuberculosis: landscape, pipeline, and unmet needs. *Discovery medicine* 13: 35–45.

University of Cape Town

## Chapter 10.

### Conclusions

*“If TB and AIDS are a snake, then the head is in South Africa while the tail is quickly moving through other African countries... And if the head of the snake is in South Africa then the teeth are in Durban”*

Dr Aaron Motsoaledi (SA Minister of Health)

World TB Day, 24 March 2011

With half a million new cases of TB annually and two out of every three of these cases also infected with HIV, South Africa (SA) is the epicenter of the dual epidemics of TB and HIV (1). TB remains the leading cause of death in individuals living with HIV infection (1) despite the availability of effective TB treatment, a well-functioning DOTs programme, and the national roll-out of anti-retrovirals. Alarming, the rates of MDR- and XDR-TB continue to increase nationally. Delayed or incorrect diagnosis leading to increases in TB transmission and greater TB-related morbidity and mortality is a major contributing factor (2). In SA less than half of all registered TB cases are smear-positive (3) and the drop-out rate amongst smear-negative culture-positive TB patients is substantial (4). Thus, ongoing research into and implementation of novel diagnostic technologies and locally relevant approaches targeting diagnostically challenging and vulnerable patient groups remains a national and global health priority.

This thesis reports studies that investigated the performance characteristics and impact of three alternative yet complementary diagnostic approaches to SN- and

SSTB at different levels of the health system. The clinical utility of the diagnostic approaches is considered within the context of high rates of routine empiric treatment. In addition, studies are reported that explore the associations of urine LAM positivity and investigate the relationship between traditional markers of TB bacillary burden (smear grade and culture time-to-positivity) and the potentially more rapid novel MTB/RIF-generated  $C_T$ -values or urine LAM concentrations.

When Dr Aaron Motsoaledi emphasized the extent of the TB and HIV problem with the aforementioned metaphor, he announced that the SA Department of Health would replace smear microscopy with the MTB/RIF assay for frontline TB diagnosis. This has ushered in a new era of molecular TB diagnosis and is certainly a major advance in addressing SNTB. Our findings in **Chapter 5** support this announcement by confirming the excellent performance of MTB/RIF in smear-positive TB, the ability of the test to diagnose more than half of all SNTB cases, and the fact that MTB/RIF can correctly detect TB cases missed by the current diagnostic reference standard of liquid culture. Furthermore, MTB/RIF-generated  $C_T$ -values were found to correlate closely with both smear grade and culture time-to-positivity, thus making it a rapid new method to quantify mycobacterial burden. This finding has important clinical and research implications. For instance, we have shown in a subsequent publication from this same patient cohort, that MTB/RIF-generated  $C_T$ -values, using a cut-point of  $>31.8$  can offer a good rule-out test for smear-positivity (NPV 80%) (5). This has relevance for public health infection control practices and contact tracing strategies. Our study also highlighted that MTB/RIF has a lower - in HIV-infected versus uninfected patients and a single test will miss ~1 in 10 culture-positive

TB cases. However, a reduced sputum sample volume can have a potential impact on test performance, emphasizing the need for improved sample acquisition methodology in SN- and SSTB.

Sputum sample induction and acquisition using the ultrasonic nebulisation of hypertonic saline is a well-established technique that is safe and effective in children and adults (6, 7), however, to date only small studies under research conditions have been conducted in adults with suspected SN- and SSTB from high HIV prevalent settings (8). Our **Chapter 3** findings confirmed SI to be safe in the largest prospective cohort of adults with suspected SN- or SSTB, but show that under programmatic conditions up to a third of SN- or SSTB suspects undergoing SI are still unable to provide a sputum specimen  $\geq 1$ ml or of adequate quality. Moreover, although culture-based TB diagnosis was  $\sim 2$ -fold higher in HIV-infected versus uninfected and inpatients versus outpatients, SI could only offer a culture-based microbiological diagnosis and hence, allow for drug susceptibility testing, in  $\sim 50\%$  of patients commencing anti-TB treatment, irrespective of patient sub-group. These findings raised concern about the utility of SI in a decentralized outpatient clinic setting. A Malawian study suggested that simple healthcare worker-provided instruction to aid sample acquisition could offer equivalent gains in culture-based diagnosis to SI (9) and thus, given the paucity of comparative data on primary care sampling strategies for adults with suspected SN- or SSTB, we designed a study to address this question.

The findings of this pragmatic randomised controlled trial presented in **Chapter 4** show that, in a primary care high HIV prevalent setting, healthcare worker-provided instruction and not induction may be the preferred initial sputum

sampling strategy. Although, perhaps unsurprisingly, SI outperformed instruction in both sampling success and culture-based diagnosis, same day diagnosis (using smear microscopy or potentially MTB/RIF) was similar between groups. This lack of difference in same-day diagnosis, combined with higher empiric treatment rates in instructed patients, meant that a similar proportion of instructed compared to induced patients initiated TB treatment during the 8-week follow-up period. Given these outcomes and the fact that healthcare-worker provided instruction is cheaper and without any side-effects, it should be the preferred initial sputum sampling strategy in primary care patients with suspected SN- or SSTB. These findings have importance for both TB clinicians and National TB programmes. To highlight these findings and raise awareness to the importance of simple healthcare-worker instruction for sputum sampling, the *Lancet Respiratory Medicine* journal organized for a podcast to accompany publication of our manuscript (<http://www.thelancet.com/lanres-audio/>, July 19<sup>th</sup> 2013).

It was surprising to note how high rates of empiric treatment attenuated the advantage of increased culture-based TB diagnosis afforded by induced sputum. This highlights how useful it is to consider diagnosis and treatment as a clinical continuum, and how valuable pragmatic and operational trials with treatment end-points are in a comprehensive evaluation of a novel TB diagnostic tool or approach. Importantly, pragmatic RCT allows for the incorporation of the rather dynamic Bayesian process of clinician decision-making and empiric treatment practice. In modeling or simple diagnostic accuracy studies, it is difficult to predict or study the effects of this dynamic clinical decision-making process. Clinical decision-making is not a frequentist process, but involves assimilating an

accumulating series of pre-test probabilities (10) (e.g. the clinical and radiological findings, together with the smear and C-reactive protein results), and weighting the overall post-test probability of disease against a threshold probability for initiating treatment. It goes contrary to the binary structure used in the standard Markov decision-tree analysis of most modeling studies to date. For this reason, empiric treatment practices may substantially reduce the anticipated benefits of novel TB diagnostic tools. More complex mathematical models that consider several contextual factors thought to modulate the potential impact of novel diagnostic tools, have been developed (11). These newer models suggest that when existing diagnostic strategies (e.g. clinical suspicion or CXR) are highly sensitive for SNTB the improvement offered by a new, more accurate diagnostic tool is significantly reduced (11).

While empiric treatment practice may affect the impact of novel diagnostic tools, is it ultimately still appropriate, or does it expose patients unnecessarily to toxic, prolonged therapy? To date, our ability to adequately answer this question has been limited by the sub-optimal performance of the TB diagnostic reference standard, liquid culture. Nevertheless, some studies are informative. A Chinese study that conducted detailed work-up of SNTB patients referred for DOTs found only 2.6% of patients to have been incorrectly diagnosed with PTB (12), suggesting that in a high TB prevalence setting the majority of empiric treatment decisions are accurate. Similarly, a Ugandan study of the WHO SNTB algorithm for HIV-infected patients (also used to make treatment decisions for the majority of empirically treated patients in the **Chapter 4** RCT) found the algorithm to have a sensitivity and specificity of 95% and 98% respectively (13). In contrast, Wilson et al. found the same algorithm to have only a sensitivity and specificity

of 80% and 44% respectively (14). Further studies with multiple diagnostic samples for TB culture/PCR and longer periods of clinical follow-up are required to adequately address this question. However, moving beyond simple accuracy and perhaps of greater clinical importance, were the findings of a recent South African study that showed early TB treatment initiation based on the 2007 WHO SNTB algorithm could reduce 8-week mortality and hospital length-of-stay in hospitalised TB HIV co-infected patients (15).

Prior to the molecular era, rapid bedside diagnosis for TB has been elusive, leaving only clinical and possibly radiological findings to guide immediate treatment decisions. However, with the development of tools such as the point-of-care urine LAM strip test, bedside diagnosis in less than half an hour is now a reality. In **Chapter 6** we show the results of the first evaluation of the urine LAM strip in hospitalised HIV-infected patients with advanced immunosuppression. Using the alternative ROC-curve selected grade-2 cut-point, we demonstrated a similar performance to the LAM ELISA with a sensitivity of ~40% in all hospitalised HIV-infected SNTB patients improving to greater than 50% in patients with a CD4  $\leq$ 200 cells/ $\mu$ l. In addition, we show that the urine LAM strip used in combination with sputum smear microscopy could potentially offer rapid diagnosis to almost three quarters of TB HIV co-infected patients.

The grade-2 and not the manufacturer's suggested grade-1 test cut-point was found to improve rule-in test utility. Inter-reader agreement and test specificity (and LR+) were optimised with the use of the grade-2 cut-point. This finding was in contrast to the only other evaluation study of the urine LAM strip test to date, conducted in an outpatient ARV-clinic setting in patients with higher median

CD4 cell counts (16). In this study, the grade-1 test cut-point showed excellent inter-reader agreement and specificity (16). We highlighted this cut-point difference in our recent correspondence to *Lancet Infectious Diseases* where we suggest that the 24 LAM-positive patients in the Lawn et al. study were strongly LAM-positive (i.e. had high OD readings) and would thus have had a grade 2 or higher band intensity (17). In contrast, a number of both TB and non-TB patients in our cohort had low ELISA ODs and as a result, faint bands were frequent. In addition, either contamination of urine samples or the in vivo presence of cross reacting fungal elements may occur more frequently in a hospital setting with bed-ridden patients than in outpatients. A number of studies have demonstrated LAM cross-reactivity with fungal elements and NTMs (18-20). Further unpublished data in different settings have recently confirmed our cut-point findings. The use of the grade-2 cut-point was agreed upon for use in future urine LAM strip test studies in an upcoming consensus expert review *BMC infectious diseases* (in press).

**Chapter 7 and 8** explore the significance of urine LAM positivity and attempt to improve our understanding of why LAM positivity is increased in HIV TB co-infection with advanced immunosuppression. The complex relationship between urine LAM concentration and compartment-specific bacillary loads is also explored. Based on these findings and that of other investigators, the increasing concentrations found in advancing immunosuppression appear to result from an interplay of the following key factors: i) higher total TB-bacillary burdens (increased LAM is associated with high smear grade and the presence of mycobacteriemia (19, 21)), and as an extension of disseminated TB; ii) Renal TB (our findings and those of Wood *et al.* support the strong association between

mycobacteruria and LAM-positivity (22, 23)); iii) the amount of free circulating LAM able to be filtered (immunocomplexing of LAM prevents filtration and may decrease with AIDS-related immunosuppression (24)); and iv) a “leaky” glomerular filtration membrane (evidenced by the association of LAM-positivity with an elevated protein/creatinine ratio). The failure of urine LAM concentration to correlate with either urine-based MTB/RIF-generated CT-values or pulmonary/extra-pulmonary culture time-to-positivity highlights the complex relationship between these factors. In addition, the relationship may be further confounded by differences between individual body-site and total bacillary loads, as well as sampling error (the inability to acquire representative samples and poor sample quality).

Urine LAM testing has only modest test performance and is not a useful biomarker to quantitate total bacillary burden. However, our **Chapter 9** findings suggest it may have important incremental clinical utility over current practice. Although urine LAM strip testing identified a similar overall proportion of TB patients to those identified by clinical and CXR suspicion alone and hence commenced on early empiric treatment urine LAM identified patients missed by routine early empiric treatment that had more advanced immunosuppression and greater disease severity. These findings are consistent with recent studies that show an association between LAM-positivity and increased mortality (25, 26).

### **Impact of the studies on the field**

This work has already impacted the ongoing battle against TB and hopefully will continue to do so. First and foremost, the work has direct relevance to practicing TB clinicians and to National TB programmes engaged in the development of guidelines and policy. Our sputum induction prospective cohort offers robust data to clinicians on the routine programmatic performance of this sample acquisition method in adults with SN- or SSTB, highlighting its limitations and indicating the patient populations that will experience the greatest benefit. Our primary care pragmatic RCT provides high quality evidence to advocate for NTPs to ensure training of healthcare-workers at TB clinics on sputum instruction and supervision methodology. With the recent introduction of low-cost sputum booths into primary care clinics in the greater Cape Town area, National TB programmes should proceed cautiously with scaling up sputum induction facilities. Focus should be on widespread uptake of the simple, effective healthcare-worker provided instruction to aid sputum sampling. The use of sputum induction should be restricted to use in children and when culture-based diagnosis is essential e.g. MDR/XDR-TB. Similarly, in the context of the national phased implementation of the MTB/RIF assay for frontline TB diagnosis, our reported study and subsequent follow-on work continues to address a number of key questions pertaining to the use of this test in routine clinical practice. In fact, our study was the first independent, published study to follow the FIND demonstration study published in September 2010 (27, 28). It raised awareness of the reduced performance of a single MTB/RIF test in HIV-infected patients, which our group further highlighted in correspondence to the *Lancet* in August 2011 (29). In addition, the suboptimal performance of MTB/RIF in the SINET study on induced sputum samples has implications for clinicians.

The urine LAM strip test work presented in this thesis has been instrumental in the evolution of this test towards use in routine clinical practice. Our reported findings are the first evaluation in the most appropriate and vulnerable target population. We describe an alternative test cut-point that may be key to ensuring maximum rule-in utility and clinical impact, and that is now advocated by other experts in the field. Finally, our consideration of urine LAM strip testing in the context of routine early empiric treatment was critical to illustrate that, despite only modest sensitivity, the test may offer incremental clinical utility over and above current routine practice. This work has created the evidence base for our recently commenced multi-centre randomised controlled trial (<http://clinicaltrials.gov/show/NCT01770730>), where we will evaluate the impact of adjunctive urine LAM strip testing as a rapid guide to early initiation of treatment in hospitalised TB HIV co-infected patients on 8-week post-discharge mortality and hospital length-of-stay. A positive finding in this RCT will likely lead to WHO endorsement and the widespread clinical roll out of urine LAM strip testing in HIV-endemic high burden settings.

This PhD thesis is also representative of the ongoing evolution in the TB diagnostics research field. There is an increasing awareness of the multitude of colliding factors, such as routine empiric treatment practices, that can ultimately impact the clinical utility of a diagnostic test or approach. However, the majority of published diagnostic research studies remain simple evaluations of accuracy, and the reporting of diagnostic research continues to be inconsistent. Sackett *et al.* proposed that diagnostic research, like research to evaluate new therapeutics, should consist of four phases, with phase I and II being simple case-controlled studies of diagnostic accuracy, phase III being prospective cohort evaluations in

the target populations, and phase IV studies being impact studies that evaluate whether patients who undergo a specific diagnostic test ultimately experience better health outcomes than similar patients who are not tested (30). Pragmatic diagnostic RCTs were described as the preferred study design to address phase IV questions. Furthermore, the STARD guidelines are now available to standardize the reporting of diagnostic studies (31, 32). This PhD thesis work targets, as appropriate, phase III and IV diagnostic research questions (30), and all reported studies follow the STARD diagnostic reporting guidelines (32). Undoubtedly, with the wide array of new diagnostic technologies available to NTPs, impact data from diagnostic RCTs will become fundamental to producing evidence-based guidelines and policy. In fact, researchers are now advocating for the consideration of contextual factors influencing the TB diagnostic test utility – such as the health-system level targeted and user skill – right from the start of test design and development (33).

### **Research priorities in SN- and SSTB**

Ongoing research is required to further investigate the diagnostic approaches presented in this thesis, as well as to investigate new technologies for the diagnosis of SN- and SSTB. With the national roll-out of MTB/RIF for frontline diagnosis, there are a number of new research priorities, many of which were outlined in Chapter 2 (2). Relevant to SN- and SSTB diagnosis and as an extension of our evaluation of MTB/RIF in urine samples, we have been investigating MTB/RIF performance in different extra-pulmonary biological fluids including cerebrospinal, pleural, pericardial and bronchoalveolar lavage. Although MTB/RIF is a fully automated test that can be performed by minimally

skilled technicians and even powered by mobile solar panels, an important unanswered question is the benefit of placing Genexpert machines directly into primary care clinics at the point-of-care. It is uncertain what impact, if any, the use of MTB/RIF in settings with high empiric treatment rates will have on treatment outcomes such as time-to-treatment and patient morbidity. Our group is currently analyzing the data from a recently completed multicentre evaluation study ([www.clinicaltrials.gov/ct2/show/NCT01554384](http://www.clinicaltrials.gov/ct2/show/NCT01554384)), and these findings should address some of these questions. Furthermore, the impact of novel diagnostic tools reducing empiric treatment practices is uncertain and large operational research studies to assess the routine impact of MTB/RIF testing on TB transmission, population-level incidence rates as well as patient-level outcomes such as TB and treatment-related morbidity and mortality are ongoing and important. In addition, a number of novel molecular TB diagnostics such as the LAMP assay (34, 35), alternative molecular platforms and other truly point-of-care technologies described in greater detail in Chapter 2 and our recent Respiriology review article are in development and under evaluation (2, 36). Many of these may be better suited to use in decentralized resource-limited settings (36).

Sputum sample acquisition will remain an important focus of diagnostic research given that the majority of novel molecular and traditional TB diagnostic technologies require sputum specimens for testing. Although SI did not impact treatment in the primary care setting, it offered superior culture-based TB diagnosis. Therefore, methods to concentrate induced sputum samples using centrifugation or glass beads/nanoparticle systems (37) to improve the sensitivity of rapid molecular assays like the MTB/RIF should be investigated

and may lead to clinical impact. Alternatively, other simple sampling methods suitable for decentralized use should be evaluated. Novel tools, such as the Lung Flute which uses the principles of wave physics to enhance sputum expectoration, show promising early results and should be investigated further (38). Likewise, the continued evaluation of non-sputum based diagnostics for SN- and SSTB is paramount.

Finally, the findings of this thesis support the continued evaluation of the urine LAM strip test for use as a diagnostic adjunct. The recently commenced LAM RCT described above will be the key determinant of whether the test should be rolled out in routine clinical practice. Additionally, ongoing research should continue to optimize the LAM assay. A more specific monoclonal LAM antibody is already under evaluation by FIND, and our group is also investigating the use of aptamers as a more sensitive method of urine LAM detection than antibody ELISA (36, 39).

### **Looking ahead**

The national roll-out of MTB/RIF in SA will reduce the burden of SNTB. Improved diagnosis and time-to-diagnosis will hopefully reduce TB transmission and incidence and even perhaps, TB-related patient morbidity and mortality, although the high routine use of empiric treatment may substantially lower the anticipated benefits. However, the problems of SSTB and lower rule-out utility of a single MTB/RIF in HIV-infected patients means that the burden and diagnostic challenges of these vulnerable patients will remain. The growing prevalence of MDR- and XDR-TB in SA will further motivate ongoing efforts to confirm a TB diagnosis to allow for drug susceptibility testing in SN- and SSTB patients. No

single test addresses all the diagnostic challenges across the diversity of resource-, healthcare- and clinical-settings. Integration of old and new technologies and continued tailoring of technology to individual, high and low burden, local and national settings is essential to optimise TB diagnosis.

The simple development of new tests will be insufficient to ensure successful scale-up or to guarantee impact for either individual patients or the global TB epidemic (40), even where an evidence-base exists. Countries vary in their abilities and interest to embrace new technologies, and it is clear that the willingness of national TB control programs and private sector clinicians to use and invest in new TB diagnostics is fundamental to the successful implementation of novel technologies. Even tools with excellent diagnostic accuracy such as MTB/RIF may have little impact unless widely and appropriately utilised.

Thus, is it important that high TB burden countries, especially the emerging economies (Brazil, Russia, India, China and South Africa) with high TB and drug resistance problems, drive the early adoption and scale-up of new technologies as well as lead the next wave of TB diagnostics innovation towards an affordable, simple POC test. In fact, only the combination of these efforts will allow advancements in TB diagnosis to significantly impact the global TB epidemic.

## References

1. Peter J, Theron G. The progression of tb diagnosis in the Hiv era: From microscopes to molecules and back to the bedside. *CME* 2011;29:404-407.
2. Peter J, van Zyl Smit R, Denkinger C, Pai M. Diagnosis of tuberculosis: State-of-the-art. In: Lange C, Migliori GB, editors. European monograph series: Tuberculosis: European Respiratory Society; 2012. p. 124-143.
3. WHO. Who report 2011: Global tuberculosis control. Geneva: World Health Organisation; 2011 No. ISBN 978 92 4 156438 0.
4. Boehme CC, Nicol MP, Nabeta P, Michael JS, Gotuzzo E, Tahirli R, Gler MT, Blakemore R, Worodria W, Gray C, Huang L, Caceres T, Mehdiyev R, Raymond L, Whitelaw A, Sagadevan K, Alexander H, Albert H, Cobelens F, Cox H, Alland D, Perkins MD. Feasibility, diagnostic accuracy, and effectiveness of decentralised use of the xpert mtb/rif test for diagnosis of tuberculosis and multidrug resistance: A multicentre implementation study. *Lancet* 2011;377:1495-1505.
5. Theron G, Pinto L, Peter J, Mishra HK, van Zyl-Smit R, Sharma SK, Dheda K. The use of an automated quantitative polymerase chain reaction (xpert mtb/rif) to predict the sputum smear status of tuberculosis patients. *Clinical infectious diseases : an official publication of the Infectious Diseases Society of America* 2012;54:384-388.
6. Zar HJ, Hanslo D, Apolles P, Swingler G, Hussey G. Induced sputum versus gastric lavage for microbiological confirmation of pulmonary tuberculosis in infants and young children: A prospective study. *Lancet* 2005;365:130-134.
7. Geldenhuys HD, Kleynhans W, Buckerfield N, Tameris M, Gonzalez Y, Mahomed H, Hussey G, Hanekom W, Hatherill M. Safety and tolerability of sputum induction in adolescents and adults with suspected pulmonary tuberculosis. *Eur J Clin Microbiol Infect Dis* 2011.

8. Hepple P, Ford N, McNerney R. Microscopy compared to culture for the diagnosis of tuberculosis in induced sputum samples: A systematic review. *Int J Tuberc Lung Dis* 2012;16:579-588.
9. Bell DJ, Dacombe R, Graham SM, Hicks A, Cohen D, Chikaonda T, French N, Molyneux ME, Zijlstra EE, Squire SB, Gordon SB. Simple measures are as effective as invasive techniques in the diagnosis of pulmonary tuberculosis in malawi. *Int J Tuberc Lung Dis* 2009;13:99-104.
10. Gill CJ, Sabin L, Schmid CH. Why clinicians are natural bayesians. *BMJ* 2005;330:1080-1083.
11. Lin HH, Dowdy D, Dye C, Murray M, Cohen T. The impact of new tuberculosis diagnostics on transmission: Why context matters. *Bull World Health Organ* 2012;90:739-747A.
12. Chang KC, Leung CC, Yew WW, Tam CM. Supervised and induced sputum among patients with smear-negative pulmonary tuberculosis. *Eur Respir J* 2008;31:1085-1090.
13. Walley J, Kunutsor S, Evans M, Thoullass J, Katabira E, Muchuro S, Matovu A. Validation in uganda of the new who diagnostic algorithm for smear-negative pulmonary tuberculosis in Hiv prevalent settings: Validating a new who diagnostic algorithm for tb. *J Acquir Immune Defic Syndr* 2011.
14. Wilson D, Mbhele L, Badri M, Morroni C, Nachega J, Chaisson RE, Maartens G. Evaluation of the world health organization algorithm for the diagnosis of Hiv-associated sputum smear-negative tuberculosis. *Int J Tuberc Lung Dis* 2011;15:919-924.
15. Holtz TH, Kabera G, Mthiyane T, Zingoni T, Nadesan S, Ross D, Allen J, Chideya S, Sunpath H, Rustomjee R. Use of a who-recommended algorithm to reduce mortality in seriously ill patients with Hiv infection and smear-negative

pulmonary tuberculosis in South Africa: An observational cohort study. *The Lancet infectious diseases* 2011;11:533-540.

16. 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. *The Lancet infectious diseases* 2012;12:201-209.

17. Peter JG, Theron G, Dheda K. Urine antigen test for diagnosis of Hiv-associated tuberculosis. *The Lancet infectious diseases* 2012;12:825; author reply 826-827.

18. Dheda K, Davids V, Lenders L, Roberts T, Meldau R, Ling D, Brunet L, van Zyl Smit R, Peter J, Green C, Badri M, Sechi L, Sharma S, Hoelscher M, Dawson R, Whitelaw A, Blackburn J, Pai M, Zumla A. Clinical utility of a commercial lam-elisa assay for tb diagnosis in Hiv-infected patients using urine and sputum samples. *PLoS One* 2010;5:e9848.

19. Boehme C, Molokova E, Minja F, Geis S, Loscher T, Maboko L, Koulchin V, Hoelscher M. Detection of mycobacterial lipoarabinomannan with an antigen-capture elisa in unprocessed urine of tanzanian patients with suspected tuberculosis. *Trans R Soc Trop Med Hyg* 2005;99:893-900.

20. Kroidl I, Clowes P, Mwakyelu J, Maboko L, Kiangi A, Rachow A, Reither K, Jung J, Saathoff E, Nsojo A, Hoelscher M. Reasons for false positive mtb-lam results - the search for causative contamination. *Int J Tuberc Lung Dis (submitted)* 2012.

21. Shah M, Martinson NA, Chaisson RE, Martin DJ, Variava E, Dorman SE. Quantitative analysis of a urine-based assay for detection of lipoarabinomannan in patients with tuberculosis. *J Clin Microbiol* 2010;48:2972-2974.

22. Peter JG, Theron G, Muchinga TE, Govender U, Dheda K. The diagnostic accuracy of urine-based xpert mtb/rif in Hiv-infected hospitalized patients who are smear-negative or sputum scarce *PLoS ONE* 2012;7:e39966.
23. Wood R, Racow K, Bekker LG, Middelkoop K, Vogt M, Kreiswirth BN, Lawn SD. Lipoarabinomannan in urine during tuberculosis treatment: Association with host and pathogen factors and mycobacteriuria. *BMC infectious diseases* 2012;12:47.
24. Wood R, Lawn SD. Challenges facing lipoarabinomannan urine antigen tests for diagnosing Hiv-associated tuberculosis. *Expert review of molecular diagnostics* 2012;12:549-551.
25. Lawn SD, Kerkhoff AD, Vogt M, Wood R. Clinical significance of lipoarabinomannan detection in urine using a low-cost point-of-care diagnostic assay for Hiv-associated tuberculosis. *AIDS* 2012;26:1635-1643.
26. Talbot E, Munseri P, Teixeira P, Matee M, Bakari M, Lahey T, von Reyn F. Test characteristics of urinary lipoarabinomannan and predictors of mortality among hospitalized Hiv-infected tuberculosis suspects in tanzania. *PLoS One* 2012;7:e32876.
27. Theron G, Peter J, van Zyl-Smit R, Mishra H, Streicher E, Murray S, Dawson R, Whitelaw A, Hoelscher M, Sharma S, Pai M, Warren R, Dheda K. Evaluation of the xpert mtb/rif assay for the diagnosis of pulmonary tuberculosis in a high Hiv prevalence setting. *American journal of respiratory and critical care medicine* 2011;184:132-140.
28. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, Allen J, Tahirli R, Blakemore R, Rustomjee R, Milovic A, Jones M, O'Brien SM, Persing DH, Ruesch-Gerdes S, Gotuzzo E, Rodrigues C, Alland D, Perkins MD. Rapid molecular

- detection of tuberculosis and rifampin resistance. *N Engl J Med* 2010;363:1005-1015.
29. Theron G, Peter J, Dheda K. Xpert mtb/rif test for tuberculosis. *Lancet* 2011;378:481; author reply 482-483.
30. Sackett DL, Haynes RB. The architecture of diagnostic research. *BMJ* 2002;324:539-541.
31. Bossuyt PM, Reitsma JB. The stard initiative. *Lancet* 2003;361:71.
32. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, Moher D, Rennie D, de Vet HC, Lijmer JG. The stard statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Ann Intern Med* 2003;138:W1-12.
33. Pai NP, Vadnais C, Denkinger C, Engel N, Pai M. Point-of-care testing for infectious diseases: Diversity, complexity, and barriers in low- and middle-income countries. *PLoS medicine* 2012;9:e1001306.
34. Boehme CC, Nabeta P, Henostroza G, Raqib R, Rahim Z, Gerhardt M, Sanga E, Hoelscher M, Notomi T, Hase T, Perkins MD. Operational feasibility of using loop-mediated isothermal amplification for diagnosis of pulmonary tuberculosis in microscopy centers of developing countries. *J Clin Microbiol* 2007;45:1936-1940.
35. Nabeta P, Gray C, Lan Nguyen Thi N, Gotuzzo E, Dheda K, Mello F, Hillemann D, Boehme C. Evaluation of a manual nucleic acid amplification test for tuberculosis detection. IUTLD meeting. Lille, France; 2011.
36. Dheda K, Ruhwald M, Theron G, Peter J, Yam WC. Point-of-care diagnosis of tuberculosis - past, present and future. *Respirology* 2012.
37. Nogueira CL, Wildner LM, Senna SG, Rovaris D, Gruner MF, Jakimiu AR, da Silva RM, Bazzo ML. Alternative sputum preparation to improve polymerase

chain reaction assay for mycobacterium tuberculosis detection. *Int J Tuberc Lung Dis* 2012;16:783-787.

38. Fujita A, Murata K, Takamori M. Novel method for sputum induction using the lung flute in patients with suspected pulmonary tuberculosis. *Respirology* 2009;14:899-902.

39. Rotherham LS, Maserumule C, Dheda K, Theron J, Khati M. Selection and application of ssdna aptamers to detect active tb from sputum samples. *PLoS One* 2012;7:e46862.

40. Cobelens F, van den Hof S, Pai M, Squire BS, Ramsay A, Kimmerling ME. Which new diagnostics for tuberculosis, and when? *The Journal of infectious diseases* 2012;205:suppl 2.

University of Cape Town

## Appendices

### 1. PhD declaration

University of Cape Town

I, Jonathan Grant Peter, do hereby declare that this thesis is based on ten journal manuscripts, one monograph chapter and one letter to the *Lancet Infectious Diseases* journal: all ten of these manuscripts have been published (Chapters 2-9). These manuscripts are presented in the style and formatting of the journal, except for the two review articles and monograph chapter which have been collated, modified, and updated to provide a systematic background to the thesis work. The manuscripts included are listed below, with a description of my contribution to each.

## **Chapter 2.**

In Chapter 2 I include modified versions of the following three reviews published during the course of my PhD studies:

i. Peter, J.G., van Zyl-Smit R.N., Denkinger, C.M., and Pai, M. "Diagnosis of tuberculosis: state-of-the-art" *European Respiratory Monograph* 2012; 58:124-143

*Professor Madhukar Pai approached me to author the above review article on TB diagnostics with the focus on active TB diagnosis in high burden settings. I reviewed all the relevant literature myself, summarised and edited the material, and drafted the manuscript in its entirety. The other authors made intellectual contributions to the overall structure of the manuscript, as well as editorial input to the final drafts.*

ii. Peter, J.G and Theron, G. "The progression of TB diagnosis in the HIV era: from microscopes to molecules and back to the bedside" *CME* October 2011 Vol. 29 No. 10  
*The South African Medical Journal editorial office approached me to author an article on current TB diagnostics in South Africa for their CME journal. I reviewed all the relevant literature myself, summarised and edited the material, and drafted the manuscript in its entirety. The other author gave intellectual and editorial input to the final drafts.*

iii. Peter, J.G., Green, C., Hoelscher, M., Mwaba, P., Zumla, A. and Dheda, K. "Urine for the diagnosis of tuberculosis: current approaches, clinical applicability, and new developments" *Curr Opin Pulm Med* 2010, 16:262-270

*I wrote this paper as part of my background reading and preparation for the section of this PhD project focused on urine-based diagnostics and in particular, urinary*

*lipoarabinomannan (LAM). I reviewed all the literature myself, summarised, and edited the material. I wrote the first draft of the entire manuscript, with the exception of the section on “urine PCR for TB diagnosis” which was authored by C. Green. The other authors made intellectual and editorial contributions to the final drafts.*

### **Chapter 3.**

Peter, J.G., Theron, G., Singh, N., Singh, A., and Dheda, K. “Sputum induction to aid the diagnosis of smear-negative or sputum-scarce TB in adults from a HIV-endemic setting” *European Respiratory Journal*. 2013. March 21. Epub ahead of print

*This manuscript and prospective cohort study was the first part of the SI component of this PhD project. The preliminary data from this project informed the selection and study design of a randomised controlled trial to evaluate SI in primary care practice. Thereafter, this study recruited for a further nine months and is the largest prospective cohort of adults with suspected SN- and SSTB collected to date to evaluate the routine performance of SI in an HIV-endemic setting. Using this large data set I was able to: i) generate accurate local diagnostic accuracy measures; and ii) compare and quantify the different performance of SI in relevant sub-groups (e.g. HIV-infected versus uninfected, outpatients versus inpatients). This data was useful for the evolution of my PhD work, as well as to inform both individual clinicians and National TB programme service needs. I was the principal investigator of this project. I initiated the work with the help of my supervisor (KD), obtained all funding, oversaw the development of the database and data entry, reviewed patient information and laboratory results, and did all statistical analysis. I wrote the full first draft of the manuscript. Other authors contributed to data capture (NS and AS), and provided conceptual and editorial input to the final drafts of the manuscript (GT and KD).*

### **Chapter 4.**

Peter, J.G., Theron, G., Pooran, A., Thomas, J., Pascoe, M. and Dheda, K. “Two sputum sample acquisition methods for smear-negative or sputum-scarce persons with suspected tuberculosis in primary-care practice: a randomised controlled trial” *Lancet Respiratory Medicine*. 2013 July 19. Epub ahead of print

*This randomised controlled trial formed one of the two central studies of my PhD. It followed on from the preliminary data of the first study, described above. This study is the*

*first study to: i) directly compare, using a pragmatic RCT, two sputum sampling strategies for sample acquisition in adults with suspected SN- and SSTB; ii) go beyond a simple evaluation of the diagnostic accuracy of sputum sampling strategies by instead evaluating the impact of different strategies on patient-important treatment outcomes. I designed this RCT with conceptual and intellectual input from my supervisor (KD). I was the principal investigator of this project. I obtained the majority of project funding (co-funding from the UCT Lung Infection and Immunity unit), oversaw the development of the database and data entry, examined and followed-up the majority of study patients, and did all data cleaning and statistical analysis. I wrote the first full draft of the manuscript. Other authors contributed to the examination of a portion of study patients (JT and MP), helped with the costing analysis (AP) and provided conceptual and editorial input to the final drafts of the manuscript (GT and KD).*

## **Chapter 5.**

Theron, G.\*, Peter, J.G.\*, van Zyl-Smit, R.N., Mishra, H., Streicher, E., Murray, S., Dawson, R., Whitelaw, A., Hoelscher, M. Sharma, S., Pai, M., Warren, R. and Dheda, K. "Evaluation of the MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting" *Am J Respir Crit Care Med.* 2011 Jul 1; 184 (1): 132-40

\*Equally contributing first authors

*This manuscript and cohort study formed two important aspects of this PhD project. Firstly, this cohort study, designed and initiated by my PhD supervisor – Professor Keertan Dheda – provided important background data on the burden and challenge of SN- and SSTB at our local Cape Town primary care clinics, and provided outpatient performance data of the urinary Lipoarabinomannan (LAM) ELISA (Dheda et al. 2010 (1)). Secondly, when the first major study of the GeneXpert MTB/RIF assay (MTB/RIF) was published in the NEJM in September 2010 (2), we decided that an evaluation of this novel assay for diagnosing smear-negative TB in our primary care setting, and the incorporation of MTB/RIF testing into our sputum induction (SI) work were critical to my PHD project. In addition, it allowed the investigation of a number of other MTB/RIF assay relevant research aims. I was involved in the final stages of recruiting this patient cohort, I examined a proportion of participants, recruited new research staff, helped to clean the database and perform statistical analysis. H. Mishra performed the majority of MTB/RIF assays in the laboratory, although both G. Theron and I received training in the laboratory performance of the assay*

and performed a proportion of the tests. G. Theron and I were equally contributing first authors, completing data cleaning and analysis and writing the manuscript. Other authors involved in the project provided conceptual intellectual input and editorial suggestions to the final drafts. All authors read and approved the manuscript. G. Theron and I managed all revisions.

## Chapter 6.

Peter, J. G., Theron, G., van Zyl-Smit, R.N., Haripersad, A., Mottay, L., Kraus, S., Binder, A., Meldau, R., Hardy, A. and Dheda, K. "Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients" *European Respiratory Journal* 2012 Nov; 40(5):1211-20

And

Peter, J.G., Theron, G. and Dheda, K. "Urine antigen test for diagnosis of HIV-associated tuberculosis" *Lancet Infectious Diseases* 2012 Nov; 12(11):825

*This study was the second central study of my PhD, and was the first evaluation of the point-of-care urine LAM strip test for TB diagnosis in hospitalised patients with suspected TB HIV co-infection and advancing immunosuppression. By examining inter-reader agreement and performing ROC-curve analysis, this study suggests a test cut-point to optimize rule-in test utility that differs from that suggested by the manufacturer. Using the data from this study, we further highlighted this test cut-point discrepancy in correspondence to Lancet Infectious Diseases (also included in my PhD submission). I was the principal investigator of this project. I obtained the majority of project funding, oversaw the development of the database and data entry, examined and followed-up the majority of study patients, and performed all the LAM strip tests. I did all data cleaning and statistical analysis. I wrote the first full draft of the manuscript. Other authors and research staff contributed to the enrollment and clinical management of study patients (AH, LM, SK and research nurses), helped me with the performance of urine LAM ELISA and strip testing (RM and AB), supervised and assisted with the statistical analysis (AH), and provided conceptual and editorial input to the final drafts of the manuscript (GT, RVZS and KD). I managed all revisions of this manuscript prior to publication.*

## **Chapter 7.**

Peter, J.G., Theron, G., Muchinga, T.E., Govender, U. and Dheda, K. "The diagnostic accuracy of urine-based MTB/RIF in HIV-infected hospitalized patients who are smear-negative or sputum scarce" 2012 *PLoS ONE* 7(7): e39966.

*This manuscript followed from the larger parent study described in Chapter 6 above. The aim of this manuscript was twofold: firstly, to determine whether the performance of urine-based MTB/RIF, with or without LAM strip (or ELISA) testing, could improve the diagnosis of SSTB in hospitalised HIV-infected patients; secondly, to investigate the association between mycobacteruria (as measured by urine-based MTB/RIF-positivity) and urine LAM positivity in HIV-infected patients. I was the principal investigator of this study and developed the main study hypothesis. I performed the data cleaning and more complicated statistical analysis of imputation and logistic regression modeling. I wrote the first full draft of the manuscript. Other authors contributed by performing the urine MTB/RIF assays (TM and UG), and provided conceptual and editorial input to the final drafts of the manuscript (GT and KD). I managed all revisions of this manuscript prior to publication. The diagnostic accuracy results and analysis for urine-based MTB/RIF and the effects of centrifugation on MTB/RIF performance formed part of a recent University of Cape Town MSc submission of a co-author Tapuwa Muchinga. I was a co-supervisor of this MSc project. Importantly, the remainder of the data (results of the logistic regression analysis) and all other sections e.g. discussion presented in this manuscript have not been nor are currently being submitted as part of any other degree.*

## **Chapter 8.**

Peter, J. G., Theron, G. and Dheda, K. "The relationship between urine LAM concentration and compartment-specific TB bacillary load in hospitalised HIV-infected patients with advanced immunosuppression" *SARJ* 2012 Volume 18 Issue 2: 43-46

*This manuscript follows on from the larger parent study described in Chapter 6, and represents further in-depth investigation into the utility of urine LAM testing. This study*

*explores the predictors of increasing urine LAM concentrations (or strip test grades) as well as the relationship between urine LAM concentration and either traditional markers of bacillary load (smear grade or TB culture time-to-positivity) or novel MTB/RIF-generated  $C_T$ -values in specific body compartments. It ties in with data presented in Chapter 3 that explores the correlation between traditional markers of bacillary burden and MTB/RIF-generated  $C_T$ -values. I performed all the data and statistical analysis for this study. I wrote the first complete draft of the manuscript. The other authors (GT and KD) provided conceptual and editorial input to the final drafts of the manuscript. I managed all revisions of this manuscript prior to publication.*

## **Chapter 9.**

Peter, J.G., Theron, G. and Dheda, K. "Can point-of-care urine LAM strip testing for tuberculosis add value to clinical decision making in hospitalised HIV-infected persons?" 2012 *PLoS ONE* 8(2): e54875

*This last study and manuscript arose from the analysis of the parent project described in Chapter 6. It represents the evolution of my work to fully evaluate the potential clinical utility of the urine LAM strip test in the context of a routine South African hospital setting where the rates of empiric TB treatment use is very high. This manuscript presents the findings of our investigation into whether urine LAM strip testing can offer any incremental utility over routine empiric treatment or a set of clinical predictors. I performed all the data and statistical analysis. I wrote the first complete draft of the manuscript. The other authors (GT and KD) provided conceptual and editorial input to the final drafts of the manuscript. I managed all revisions of this manuscript prior to publication.*

I confirm that no part of this thesis (other than the referenced sections clearly indicated in Chapter 7) has been submitted in the past, or is being, or is to be submitted for a degree in this or any other university. I hereby grant the University of Cape Town free license to reproduce this thesis in whole or part for the purposes of research or teaching.

This thesis is presented for examination in fulfillment of the requirements for the degree of Doctor of Philosophy in Medicine.

Signed,

A handwritten signature in black ink, appearing to read 'J. Peter', with the name 'Jonathan Grant Peter' printed in a smaller font below it. The signature is written over a faint, illegible background.

Jonathan Grant Peter

10 August 2013

As the senior author and supervisor of all the above original manuscripts, and on behalf of all co-authors, I confirm the above authors' contributions to be accurate.

A handwritten signature in black ink, appearing to read 'K Dheda', with a long horizontal stroke underneath. The signature is written over a faint, illegible background.

Professor Keertan Dheda

10 August 2013

## 2. Chapter 4 supplementary material

Peter, J.G., Theron, G., Pooran, A., Thomas, J., Pascoe, M. and Dheda, K. "Two sputum sample acquisition methods for smear-negative or sputum-scarce persons with suspected tuberculosis in primary-care practice: a randomised controlled trial" *Lancet Respiratory Medicine*. 2013 July 19. Epub ahead of print

University of Cape Town

## **Online supplementary material**

### **Evaluation of two sputum sample acquisition methods for the diagnosis of suspected tuberculosis in smear-negative or sputum-scarce persons in primary care practice: a randomised controlled trial**

Authors: Jonathan G. Peter<sup>1</sup>, Grant Theron<sup>1</sup>, Anil Pooran, Johnson Thomas, Mellissa Pascoe, Keertan Dheda<sup>1,2,3†</sup>

<sup>1</sup>Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, Faculty of Health Sciences, University of Cape Town, South Africa.

<sup>2</sup>Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, South Africa.

<sup>3</sup>Department of Infection, University College London Medical School, United Kingdom.

†Corresponding author: Keertan Dheda, +27214047651 (fax), +27214046029 (telephone), email: [Keertan.Dheda@uct.ac.za](mailto:Keertan.Dheda@uct.ac.za)

University of Cape Town

## Supplementary methods section

### Procedures

All study research nurses (a total of 10 during the course of the study) underwent at least two days of procedural training at the Grootte Schuur Hospital Respiratory clinic. Research nurses were rotated between the three primary clinic enrolment sites. During the study, to check protocol adherence either the principal investigator or another member of the study team conducted frequent unannounced checks.

#### *Healthcare worker-provided instruction*

Patients allocated to receive healthcare worker-provided instruction were individually trained by the study nurse in their preferred language from English, Afrikaans, Xhosa, Sotho and Zulu. The guidance began by emphasising the importance of providing a sputum rather than a salivary sample. The technique to produce a good sputum specimen, namely to take three deep breaths followed by a deep cough (1), was then explained and demonstrated. Patients were then given two sputum containers and told to fill at least a quarter of each container if possible. Patients were instructed to first fill one sputum container to an adequate volume (~1ml) before filling the second.

#### *Healthcare worker-provided instruction script*

(To be told to the patient in their preferred language as above)

1. Describe to the patients the difference between saliva and sputum samples (saliva: clear, watery fluid versus sputum: thick, viscous/mucoid, often not white)
2. Give patient two sputum jars and ask patient to provide a sputum sample of at least 1ml in each jar (indicate that 1ml of fluid will cover the bottom of the container completely)
3. Describe and demonstrate how to best produce sputum, namely to take 3 deep breaths and then give a deep cough
4. Observe patient during sputum production (wearing an N95 mask) and repeat instruction no. 3 if necessary

#### *Sputum induction*

Patients allocated to sputum induction were asked to rinse their mouths with water and stand in an outdoor, open-air ventilated booth (Picture S1). Approximately 20 ml of sterile 5% hypertonic saline (Sabax, South Africa) was delivered via a Wilson's 402A ultrasonic nebuliser (Medimark, South Africa, <http://www.medimark.co.za/>, contact tel: +27215101995). Flow rate was adjusted until a fine white mist emerged from the mask. Patients breathed through the nebuliser for up to 20 minutes or until a satisfactory sample was produced. No pro-expectorating maneuvers, such as chest percussion, were performed to assist sputum production. Patients were given two sputum containers and told to fill at least a quarter of each container if possible. Patients were instructed to first fill one sputum container to an adequate volume (~1ml) before filling the second. Induction was immediately terminated if side effects such as dyspnoea, shortness of breath, chest pain or nausea were reported by patients.

**Picture S1: Langa clinic outdoor, open-air, low-cost sputum induction booth**



## Laboratory methods

All patients were asked to provide two spot sputum samples during the intervention (Figure 1, main manuscript). Samples of  $\geq 1$  ml, irrespective of visual quality, were sent for processing at the National Health Laboratory services located centrally. When patients provided two samples they were randomly labelled sputum 1 or sputum 2. Sputum specimens were transported by courier to the laboratory and results for smear microscopy were available within 24 hours and not same-day. Sputum 1 was processed with N-acetyl-L-cysteine and sodium hydroxide (NALC-NaOH), centrifuged, and resuspended in 1.5ml phosphate buffer. Thereafter, the sample was subjected to Auramine-O staining and fluorescence microscopy and 0.5ml of the sediment was inoculated into a Mycobacterial Growth Indicator Tube (MGIT; Becton Dickinson Diagnostics, USA) and incubated for  $\leq 8$  weeks. All postivity culture results were phoned through to the research nursing staff within 2 days of becoming positive. Sputum 2 was unprocessed and frozen at  $-20^{\circ}\text{C}$  within 6 hours of acquisition by our research laboratory. After study completion, the available second sputa were thawed and used for Xpert MTB/RIF testing (2). If patients provided only a single sputum sample, this was processed as sputum 1.

University of Cape Town

## Supplementary results section

### Characteristics of patients giving one versus two sputum samples

**Table S1. Baseline demographic and clinical characteristics of patients stratified by number of sputum samples produced.**

Demographic and clinical characteristic(s)	All (N=481)	Patients producing only a single sample (N=254)	Patients producing two samples (N=227)
<b>Demographic characteristics</b>			
Median age (years, IQR)	39 (30-49)	38 (30-48)	39 (31-50)
Male sex (%)	262 (55)	143 (56)	119 (52)
HIV-infected (%)	171 (36)	93 (37)	78 (34)
Median CD4 cell count (cells/ml, IQR)	242 (146-358)	219 (129-325)	257 (167-385)
Current ARVs (%)	37 (22)	15 (16)	22 (28)
History of TB (%)	180 (37)	92 (36)	88 (39)
<b>Diagnostic categorization at enrolment</b>			
2 × sputum smear-negative (%)	244 (51)	137 (54)	107 (47)
Unable to produce sputum prior to enrolment (%)	237 (49)	117 (46)	120 (53)
<b>Symptoms, signs and radiological features at enrolment</b>			
Cough > 2 weeks (%)	430 (90)	218 (86)**	212 (94)**
Productive cough (%)	311 (65)	140 (56)***	171 (76)***
Night sweats (%)	344 (72)	181 (72)	163 (72)
Weight loss (%)	335 (70)	176 (70)	159 (70)
Appetite loss (%)	253 (54)	136 (55)	117 (54)
Weight (median kg, IQR)	62 (54-72)	62 (54-71)	62 (54-74)
CXR compatible with TB	179 (37)	103 (41)	76 (35)

†P-values indicate comparison between study groups for demographic or clinical characteristic (marked with \*); \*\*p<0.01, \*\*\*p<0.001

HCW: Health care worker; IQR: interquartile range; ARV: anti-retroviral therapy; CXR: chest x-ray; Tuberculosis, TB

## Component costs used for costing sampling procedures

**Table S2: Component costs of each sputum sampling procedures.**

Healthcare-worker provided instruction and supervision of self expectorated sputum		Sputum induction using an ultrasonic nebulizer	
Cost Components	Unit cost (\$US)	Cost Component	Unit cost (\$US)
Consumables		Consumables	
Sputum jars	\$0.46	nebulizer tubing, masks, syringes and needles, saline, sputum jars	\$5.05
<b>Capital costs</b>		<b>Capital costs</b>	
Clinic overheads	\$0.51	Clinic overheads	\$0.51
		annualized cost per use of nebulizer	\$0.01
<b>Staff costs</b>		<b>Staff costs</b>	
One day staff training per sampling cost (every six months)	\$0.01	Two day staff training per sampling cost (every six months)	\$0.01
Patient supervision by clinic nurse (10 minutes)	\$1.16	Patient supervision by clinic nurse (20 minutes)	\$2.31
<b>TOTAL</b>	<b>\$2.14</b>	<b>TOTAL</b>	<b>\$7.88</b>

Consumable costs are based on current South African market prices. Capital costs were annualized at a discount rate of 3%. The expected life of buildings was estimated to be 10 years. The expected life of an ultrasonic nebulizer was estimated to be 2 years. Staff costs for each sampling strategy were calculated based on Provincial Government of Western Cape salary scales and the estimated time required for staff training and patient supervision.

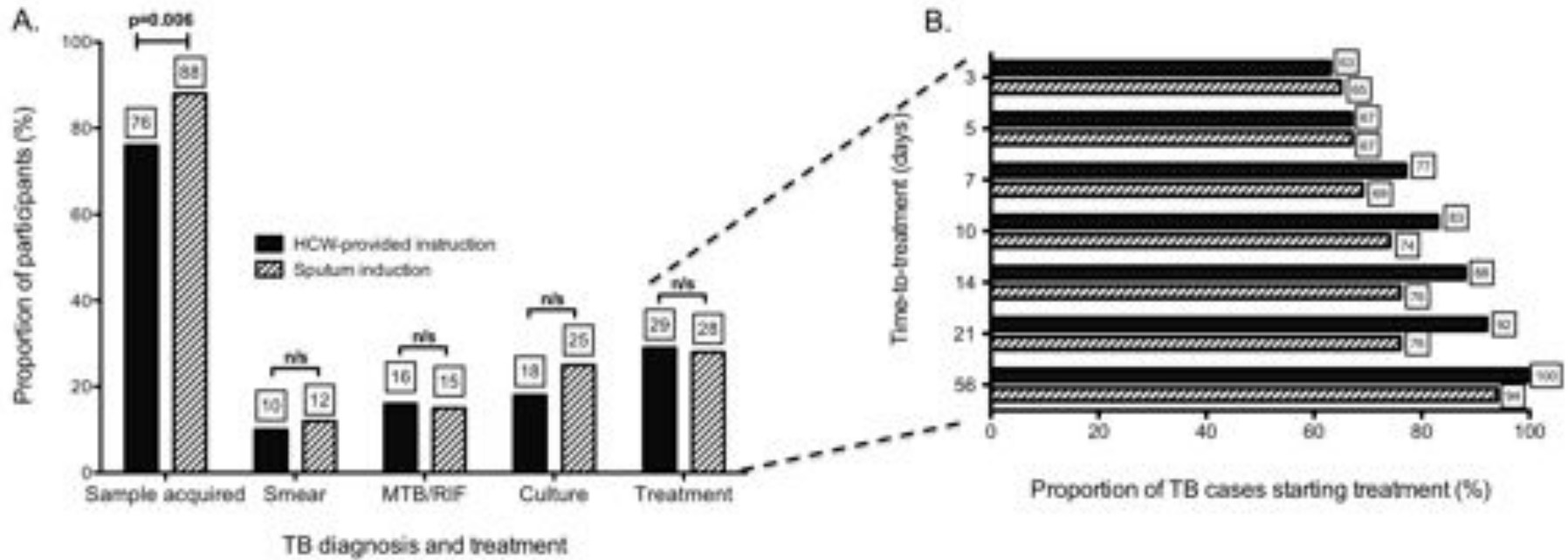
Key diagnostic and treatment outcomes if only patients able to successfully provide a sputum sample for diagnostic testing (irrespective of sampling strategy) are included (N=402; HCW-provided instruction=164, Sputum induction=238)

Figure S1: The comparative proportions of patients who received healthcare worker-provided instruction or sputum induction for key diagnostic and treatment outcomes, including: (A) Smear-microscopy, Xpert MTB/RIF and culture-based diagnostic yields, and patients given treatment; and (B) proportions of patients initiating treatment at specific time-points from enrollment.

Treatment characteristic	HCW-provided instruction (N=36)	Sputum induction (N=68)	P-value
Median time-to-treatment (days)	6 (2-14)	7 (3-28)	0.2

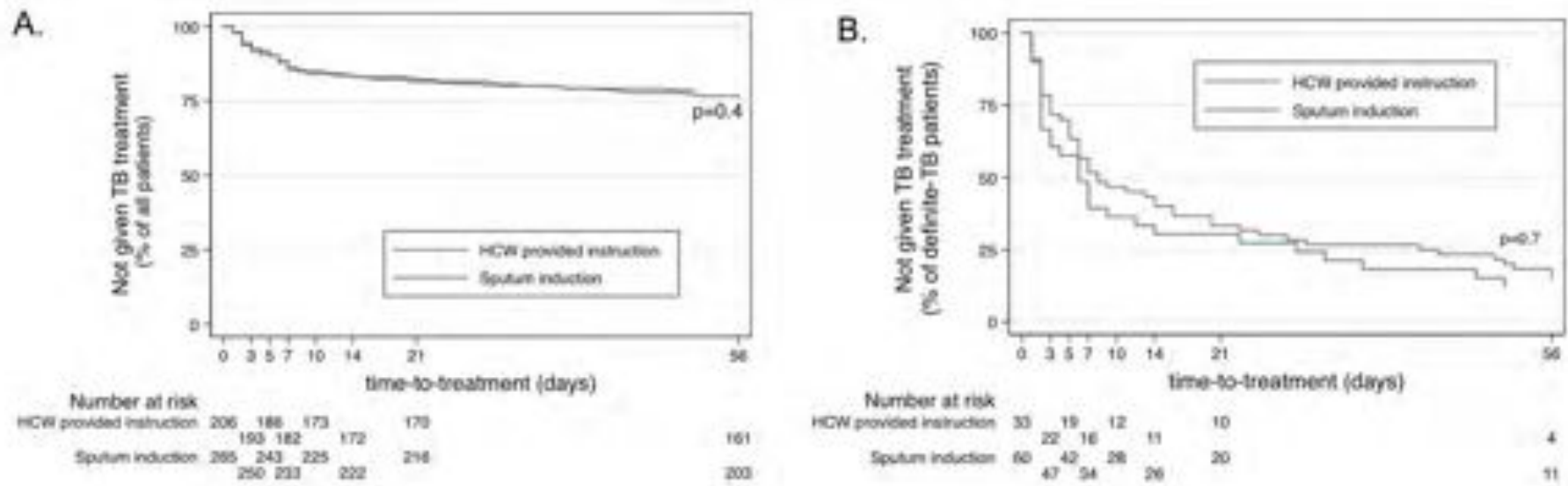
Key diagnostic and treatment outcomes using a random selection of 200 patients from each study arm

Figure S2: The comparative proportions of patients who received healthcare worker-provided instruction or sputum induction for key diagnostic and treatment outcomes, including: (A) Smear-microscopy, Xpert MTB/RIF and culture-based diagnostic yields, and patients given treatment; and (B) proportions of patients initiating treatment at specific time-points from enrollment.



Alternative survival-type analysis for time-to-treatment initiation

Figure S3: Kaplan Meier estimates comparing the time-to-treatment initiation between study arms in A) All study patients (N=481) and B) Only definite-TB patients (N=98).



P-values are for a logrank test comparison between study arm time-to-treatment curves  
 X-axis labelling intervals selected to correspond with the time-points used in the primary analysis  
 HCW: Healthcare worker

## References

1. Khan MS, Dar O, Sismanidis C, Shah K, Godfrey-Faussett P. Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial. *Lancet*. 2007 Jun 9;369(9577):1955-60. PubMed PMID: 17560448. Epub 2007/06/15. eng.
2. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F, et al. Rapid molecular detection of tuberculosis and rifampin resistance. *N Engl J Med*. 2010 Sep 9;363(11):1005-15. PubMed PMID: 20825313. Pubmed Central PMCID: 2947799. Epub 2010/09/10. eng.

University of Cape Town

### 3. Chapter 5 supplementary material

Theron, G.\*, Peter, J.G.\*, van Zyl-Smit, R.N., Mishra, H., Streicher, E., Murray, S., Dawson, R., Whitelaw, A., Hoelscher, M. Sharma, S., Pai, M., Warren, R. and Dheda, K. "Evaluation of the MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting" *Am J Respir Crit Care Med.* 2011 Jul 1; 184 (1): 132-40

\*Equally contributing first authors

University of Cape Town

## **ONLINE DATA SUPPLEMENT**

### **Evaluation of the Xpert<sup>®</sup> MTB/RIF assay for the diagnosis of pulmonary tuberculosis in a high HIV prevalence setting**

Grant Theron<sup>1\*</sup>, Jonny Peter<sup>1\*</sup>, Richard van Zyl-Smit<sup>1</sup>, Hridesh Mishra<sup>2</sup>, Elizabeth Streicher<sup>3</sup>, Samuel Murray<sup>1</sup>, Rodney Dawson<sup>1</sup>, Andrew Whitelaw<sup>4</sup>, Michael Hoelscher<sup>5</sup>, Surendra Sharma<sup>2</sup>, Madhukar Pai<sup>6</sup>, Robin Warren<sup>3</sup>, Keertan Dheda<sup>1, 7, 8†</sup>

<sup>1</sup>Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, University of Cape Town, South Africa.

<sup>2</sup>Department of Medicine, All India Institute of Medical Sciences, India.

<sup>3</sup>DST/NRF Centre of Excellence for Biomedical TB Research/MRC Centre for Molecular and Cellular Biology, Stellenbosch University, South Africa.

<sup>4</sup>Division of Medical Microbiology, University of Cape Town, Cape Town, South Africa.

<sup>5</sup>Department for Infectious Diseases and Tropical Medicine, Klinikum of the University of Munich, Munich, Germany.

<sup>6</sup>Department of Epidemiology & Biostatistics, McGill University, Montreal, Canada.

<sup>7</sup>Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, South Africa.

<sup>8</sup>Department of Infection, University College London Medical School, United Kingdom.

Requests for reprints should be addressed to Keertan Dheda, H47 Old Main Building, Groote Schuur Hospital, Observatory, 7925, South Africa or keertan.dheda@uct.ac.za

\*The first two authors contributed equally.

†Corresponding author: keertan.dheda@uct.ac.za, +27214047651 (fax), +27214046509 (telephone)

**Table S1.** Positive and negative likelihood ratios of Xpert<sup>®</sup> MTB/RIF for the detection of *M. tuberculosis* compared to smear microscopy, and stratified by HIV status\*.

	All patients		HIV uninfected patients		HIV infected patients	
	n = 480		n = 286		n = 130	
	Positive LR (95% CI)	Negative LR (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)
Sputum smear	226 (31.5 - 1621)	0.33 (0.32 - 0.35)	N/A*	0.27 (0.26 - 0.29) <sup>‡</sup>	42.0 (5.4 - 324.7)	0.51 (0.47 - 0.55) <sup>‡</sup>
Xpert <sup>®</sup> MTB/RIF	14.1 (12.61 - 15.65) <sup>†</sup>	0.23 (0.21 - 0.24) <sup>†</sup>	18.8 (15.0 - 23.5)	0.18 (0.16 - 0.21) <sup>†</sup>	8.4 (6.1 - 11.4)	0.33 (0.29 - 0.38) <sup>†‡</sup>
Sputum smear and/or Xpert <sup>®</sup> MTB/RIF	13.9 (12.60 - 15.44) <sup>†</sup>	0.19 (0.17 - 0.20) <sup>†</sup>	19.4 (15.5 - 24.2)	0.15 (0.13 - 0.18) <sup>†</sup>	7.8 (6.0 - 10.1)	0.29 (0.24 - 0.34) <sup>†‡</sup>

\* We are unable to calculate LRs in the Xpert<sup>®</sup> MTB/RIF-positive smear -ve, culture +ve group (as shown in Table 2A), given that the number of true negative cases is zero. Similarly, we are unable to calculate the LR where indicated due to either (or both) of the number of true negatives or false negatives being zero.

<sup>†</sup> Indicates significantly different LRs (non-overlapping CI) (Xpert<sup>®</sup> MTB/RIF vs. sputum smear or sputum smear vs. a combination of both) on a per patient subgroup basis

<sup>‡</sup> Indicates assay specific significantly different LRs (non-overlapping CI) between patient subgroups (HIV-infected patients vs. HIV-uninfected patients).

**Table S2.** Positive and negative likelihood ratios of Xpert<sup>®</sup> MTB/RIF for the detection of *M. tuberculosis* compared to smear microscopy in HIV-infected persons, and stratified by CD4 count.\*

	HIV infected patients		HIV infected patients with CD4 count $\geq$ 200 cells/ml		HIV infected patients with CD4 count < 200 cells/ml	
	n = 130		n = 57		n = 66	
	Positive LR (95% CI)	Negative LR (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)	Positive LR (95% CI)	Negative LR (95% CI)
Sputum smear	42 (5.4 - 324.7)	0.51 (0.46 - 0.55)	22.29 (2.86 - 173.6)	0.39 (0.31 - 0.50)	N/A*	0.61 (0.53 - 0.70)
Xpert <sup>®</sup> MTB/RIF	8.35 (6.14 - 11.35)	0.33 (0.29 - 0.38) †‡	27.43 (3.72 - 202.3)	0.24 (0.17 - 0.36)	9.35 (4.54 - 19.26)	0.37 (0.29 - 0.48) †‡
Sputum smear and/or Xpert <sup>®</sup> MTB/RIF	7.76 (6.0 - 10.12)	0.29 (0.24 - 0.34) <sup>†</sup>	14.17 (5.185 - 38.78)	0.20 (0.12 - 0.33)	9.97 (4.92 - 20.22)	0.33 (0.25 - 0.43) <sup>†</sup>

\* We are unable to calculate LRs in the Xpert<sup>®</sup> MTB/RIF-positive smear -ve, culture +ve group (as shown in Table 2A), given that the number of true negative cases is zero. Similarly, we are unable to calculate the LR where indicated due to either (or both) of the number of true negatives or false negatives being zero.

<sup>†</sup> Indicates significantly different LRs (non-overlapping CI) (Xpert<sup>®</sup> MTB/RIF vs. sputum smear or sputum smear vs. a combination of both) on a per patient subgroup basis

<sup>‡</sup> Indicates assay specific significantly different LRs (non-overlapping CI) between patient subgroups (HIV-infected vs. HIV-uninfected patients or HIV infected patients with CD4 count < 200 cells/ml vs. HIV-uninfected patients).

**Table S3.** Xpert<sup>®</sup> MTB/RIF positive and negative likelihood ratios in patients with definite TB (all culture positive cases) versus a group containing culture-positive and highly likely TB cases (as defined in the footnote; p values indicate a comparison between groups). The same analysis for patients with smear-negative TB is also shown.

		Positive LR (95% CI)	Negative LR (95% CI)
All TB cases	Definite TB (culture positive)	14.05 (12.61 - 15.65)	0.23 (0.21 - 0.24)
	Definite TB + highly likely group*	87.09(45.15 - 168)	0.19 (0.18 - 0.1)
Smear-negative TB	Definite TB (culture positive)	8.33 (6.79 - 10.22)	0.56 (0.5209 - 0.6098)
	Definite TB + highly likely group*	64.74 (32.56 - 128.7)	0.40 (0.37 - 0.43)

\*16 culture negative, Xpert<sup>®</sup> MTB/RIF-positive patients are included here (5 patients were found to be culture-positive by a second sputum obtained within 2 weeks from enrolment, 5 had *M. tuberculosis* DNA in their sputum by sequencing, and 6 patients had typical radiological evidence of active TB). All Xpert<sup>®</sup> MTB/RIF amplicons were confirmed to contain *M. tuberculosis* DNA.

#### 4. Chapter 6 supplementary material

Peter, J. G., Theron, G., van Zyl-Smit, R.N., Haripersad, A., Mottay, L., Kraus, S., Binder, A., Meldau, R., Hardy, A. and Dheda, K. "Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients" *European Respiratory Journal* 2012 Nov; 40(5):1211-20

University of Cape Town

## **Online supplementary material**

### **Diagnostic accuracy of a urine LAM strip-test for TB detection in HIV-infected hospitalised patients**

Authors: Jonathan G. Peter<sup>1</sup>, Grant Theron<sup>1</sup>, Richard van Zyl-Smit<sup>1</sup>, Asheen Haripersad<sup>1</sup>, Lynelle Mottay<sup>1</sup>, Sarah Kraus<sup>1</sup>, Anke Binder<sup>1</sup>, Richard Meldau<sup>1</sup>, Anneli Hardy<sup>2</sup> and Keertan Dheda<sup>1,3,4†</sup>

<sup>1</sup>Lung Infection and Immunity Unit, Division of Pulmonology & UCT Lung Institute, Department of Medicine, University of Cape Town, South Africa.

<sup>2</sup>Department of Statistical Sciences, University of Cape Town, South Africa.

<sup>3</sup>Institute of Infectious Diseases and Molecular Medicine, University of Cape Town, South Africa.

<sup>4</sup>Department of Infection, University College London Medical School, United Kingdom.

†Corresponding author: Keertan Dheda, +27214047651 (fax), +27214046029 (telephone), email: Keertan.Dheda@uct.ac.za

## **Supplementary methods section**

### **LAM methodology**

All patients were required to give a spot urine sample (10-30ml) collected in a sterile container as soon as possible after recruitment. A urine dipstick test (*UriCHECK 9*, RapiMed Diagnostics, South Africa) was immediately performed to assess for protein, blood and leucocytes. Urine was stored at -20°C for later batched testing. The LAM strip test (a single manufacturing lot #101102) was performed on unprocessed urine according to manufacturer's instructions. Briefly, urine was thawed, mixed and 60µl pipetted onto the lateral flow strip loading bay. After 25 minutes, two readers in ambient laboratory lighting conditions, and blinded to the clinical patient details and TB status, independently evaluated the LAM strips for all study patients via the following procedure: after confirming test validity by identifying the presence of a band in the positive control window, the intensity of the band (if any) in the patient window was read using a manufacturer-provided visual reference scale card (graded 0 – 5 depending on band intensity; see Figure 1 in manuscript). Using the manufacturer-recommended grade 1 cut-point, a band of visual intensity  $\geq$  grade 1 in the patient window was classified test 'positive' while only the complete absence of a band (grade 0) in the patient window was classified test 'negative' (see Figure 1). Accuracy was assessed at various alternative cut-points to select one for optimal rule-in value. For example, if the grade 2 cut-point was selected the complete absence of band (grade 0) as well as a faint band (grade 1) was classified test 'negative'. The test was reported as indeterminate if a broken/ incomplete band was seen in the patient window. A test was reported as failed if no control band was identified. Each reader graded the LAM strips twice, the second reading occurring 5-10 minutes after the initial evaluation and readers were blinded to the result of their initial grading. After TB status was determined, to further evaluate inter-observer

reliability at the grade 1 and 2 cut-point, a group of 4 independent readers (different from the initial readers) graded a second urine LAM strip test (also from manufacturing lot #101102) for each patient in the definite- and non-TB groups only. Additional to 2 readers independently grading one LAM strip for the entire study cohort, Concurrently, the Clearview TB ELISA was performed on all samples as previously described, by a technician blinded to patient TB diagnostic status (22).

University of Cape Town

## Supplementary results section

*Supplementary results to show positive likelihood ratios for analysis 1 and 2*

### *A. Overall*

Table A1. Performance outcomes of smear microscopy, LAM ELISA, and LAM strip test (using grade 1 and 2 cut-points) in HIV-infected patients for the analysis 1 (microbiological reference standard)<sup>†</sup>.

Diagnostic test(s)	Group 1 patients (N=242)		Group 2 non-TB control patients (N=88)	Positive likelihood ratio <sup>§</sup> (using group 1 patient specificity) (95% CI)	Positive likelihood ratio <sup>§</sup> (using group 2 patient specificity) (95% CI)
	Sensitivity (95% CI)	Specificity (95% CI)	Specificity (95% CI)		
<b>Rapid smear microscopy<sup>‡</sup></b>	56 <sup>#2</sup> (47-65) 65/116	90 <sup>#4</sup> (84-95) 114/126	N/R	5.9 (4.9-7.0)	N/C
<b>LAM ELISA<sup>&amp;</sup></b>	59 (50-68) 68/115	80 <sup>#4 #5 *1</sup> (72-86) 98/123	99 <sup>*1</sup> (94-100) 87/88	2.9 (2.6-3.2)	52.0 (7.2-376.8)
<b>LAM strip test<sup>&amp;</sup> (grade 1 cut-point)</b>	66 <sup>#1</sup> (57-74) 77/116	66 <sup>#5 *2</sup> (57-73) 82/125	90 <sup>#6 *2</sup> (82-95) 79/88	1.9 (1.8-2.0)	6.5 (5.2-8.2)
<b>LAM strip test<sup>&amp;</sup> (grade 2 cut-point)</b>	50 <sup>#1 #3</sup> (41-59) 58/116	75 <sup>*3</sup> (67-82) 94/125	99 <sup>#6 *3</sup> (94-100) 87/88	2.0 (1.8-2.2)	44 (6.0-323.1)
<b>LAM strip test (grade 2 cut-point) in smear-negative/sputum scarce <i>M.tb</i> culture positive patients</b>	39 (28-52) 22/56	N/A	N/A	N/A	N/A
<b>Combined sputum smear microscopy and LAM strip test (grade 2 cut-point)</b>	71 <sup>#2 #3</sup> (62-78) 82/116	73 <sup>*4</sup> (65-80) 92/126	99 <sup>*4</sup> (94-100) 87/88	2.6 (2.5-2.8)	62.2 (8.7-446)

<sup>†</sup>242 HIV-infected group 1 patients with  $\geq 1$  valid TB culture result were available for this primary analysis. 116 and 126 *M.tb* culture positive and negative patients were used for group 1 sensitivity and specificity calculations, respectively. 88 Group 2 non-TB control patients were used to calculate a second specificity.

§2 different positive likelihood ratios are calculated for each test using the specificity calculated for group 1 and 2 patients respectively

‡Rapid smear microscopy includes both sputum and non-sputum samples (e.g. fine-needle aspirate and cerebrospinal fluid samples) attainable within 24 hours of hospital admission.

&1 and 3 *M.tb* culture positive and negative patients had a missing LAM ELISA result. 1 *M.tb* culture negative patient had missing urine LAM strip test.

#Indicates  $p < 0.05$  for a comparison of the specific measures of diagnostic accuracy (sensitivity or specificity) between different tests (LAM ELISA vs. LAM strip test) or combinations thereof, or between different cut-points of the urine LAM strip test (grade-1 vs. grade-2 cut-point); specific p-value: #<sup>1</sup> $p=0.01$ ; #<sup>2</sup> $p=0.02$ ; #<sup>3</sup> $p=0.001$ ; #<sup>4</sup> $p=0.02$ ; #<sup>5</sup> $p=0.01$ ; #<sup>6</sup> $p=0.009$ ; non-significant p-values not shown)

\* Indicates  $p < 0.05$  for a comparison of differences in diagnostic accuracy measures between analyses and patient groups for a specific test (LAM ELISA or LAM strip test); specific p-values: \*<sup>1</sup> $p < 0.001$ ; \*<sup>2</sup> $p < 0.001$ ; \*<sup>3</sup> $p < 0.001$ ; \*<sup>4</sup> $p < 0.001$ )

N/A: not applicable; N/C: unable to be calculated, N/R: no results

University of Cape Town

Table A2 Performance outcomes of smear microscopy, LAM ELISA, and LAM strip test (using grade 1 and 2 cut-points) in HIV-infected patients for analysis 2 (composite reference standard)<sup>†</sup>.

Diagnostic test(s)	Group 1 and 2 patients combined		
	N=302		
	Sensitivity (95% CI)	Specificity <sup>§</sup> (95% CI)	Positive likelihood ratio (95% CI)
Rapid smear microscopy <sup>‡</sup>	44 <sup>#1</sup> (37-51) 82/187	99 (95-100) 114/115	50.4 (6.9-369.1)
LAM ELISA <sup>&amp;</sup>	51 (44-58) 93/183	98 (94-100) 111/113	28.7 (10.6-78.1)
LAM strip test <sup>&amp;</sup> (grade 1 cut-point)	60 (53-67) 112/187	91 <sup>#3</sup> (85-95) 105/115	6.9 (5.6-8.5)
LAM strip test <sup>&amp;</sup> (grade 2 cut-point)	45 <sup>#2</sup> (39-53) 85/187	97 <sup>#3</sup> (94-100) 113/115	26.1 (9.5-71.6)
LAM strip test (grade 2 cut-point) in smear-negative/sputum scarce <i>M.tb</i> culture positive patients	38 (29-48) 36/95	N/A	N/A
Combined sputum smear microscopy and LAM strip test (grade 2 cut-point)	63 <sup>#1 #2</sup> (56-70) 118/187	97 (93-99) 112/115	24.2 (12.5-46.9)

<sup>†</sup>302 definite-, probable- and non-TB HIV-infected patients with known HIV status were available for this secondary analysis. 116 definite-TB and 71 probable-TB were used for sensitivity calculations and 115 non-TB (27 group 1 and 88 group 2) were used for specificity calculations. The unclassifiable-TB patient group was excluded from this analysis.

<sup>§</sup>Tabulated LAM strip test specificities were calculated for HIV-infected patients using group 1 and 2 non-TB patients combined. Specificity results for HIV-infected group 1 and 2 non-TB patients separately, using the grade 1 cut-point, were 96% (82-99) and 90% (82-95) respectively and, using the grade 2 cut-point, were 96% (82-99) and 99% (94-100) respectively. No significant difference in specificity was noted between HIV-infected groups 1 and 2 non-TB patients.

<sup>‡</sup>Rapid smear microscopy includes both sputum and non-sputum samples (e.g. fine-needle aspirate and cerebrospinal fluid samples) attainable within 24 hours of hospital admission.

<sup>&</sup>6 HIV infected patients (4 definite-TB and 2 non-TB) had a missing LAM ELISA result

#Indicates  $p < 0.05$  for a comparison of the specific measures of diagnostic accuracy (sensitivity or specificity) between different tests (LAM ELISA vs. LAM strip test) or combinations thereof, or between different cut-points of the urine LAM strip test (grade-1 vs. grade-2 cut-point); specific p-value: #<sup>1</sup> $p < 0.001$ ; #<sup>2</sup> $p < 0.001$ ; #<sup>3</sup> $p = 0.02$ ; non-significant p-values not shown)

N/A: not applicable; N/C: unable to be calculated

University of Cape Town

## B. Stratified by CD4 cell count

Table B1. Performance outcomes of smear microscopy, LAM ELISA, and LAM strip test (using grade 1 and 2 cut-points) in HIV-infected patients stratified by CD4 count (>200, ≤200 and ≤100 cells/ml) in analysis 1<sup>†</sup>.

Diagnostic test(s)	HIV-infected patients with CD4 count >200 cells/ml			HIV-infected patients with CD4 count ≤200 cells/ml			HIV-infected patients with CD4 count ≤100 cells/ml		
	N=55			N=173			N=120		
	Sensitivity (95% CI)	Group 1 Specificity (95% CI)	Group 2 Specificity (95% CI)	Sensitivity (95% CI)	Group 1 Specificity (95% CI)	Group 2 Specificity (95% CI)	Sensitivity (95% CI)	Group 1 Specificity (95% CI)	Group 2 Specificity (95% CI)
<b>Rapid smear microscopy<sup>‡</sup></b>	58 (39-75) 15/26	90 (79-100) 26/29	N/R	56 <sup>#2</sup> (45-66) 45/81	91 <sup>#4 #5</sup> (86-97) 84/92	N/R	54 <sup>#6</sup> (42-67) 31/57	90 <sup>#7 #8</sup> (83-98) 57/63	N/R
<b>LAM ELISA<sup>&amp;</sup></b>	27 <sup>*1</sup> (14-46) 7/26	89 (77-100) 24/27	100 (93-100) 53/53	70 <sup>*1</sup> (59-79) 56/80	76 <sup>#4 *5</sup> (67-85) 69/91	97 <sup>*5</sup> (90-100) 28/29	75 (62-85) 42/56	69 <sup>#7</sup> (59-81) 43/62	93 (79-100) 13/14
<b>LAM strip test (grade 1 cut-point)</b>	54 (36-71) 14/26	76 (60-91) 22/29	91 <sup>#1</sup> (83-99) 49/54	72 (61-80) 58/81	61 (51-71) 56/92	87 (75-99) 26/30	77 (65-86) 44/57	65 (53-77) 41/63	87 (69-100) 13/15
<b>LAM strip test (grade 2 cut-point)</b>	27 <sup>*2</sup> (14-46) 7/26	83 <sup>*3</sup> (69-97) 24/29	100 <sup>*3 #1</sup> (93-100) 54/54	58 <sup>*2 #3</sup> (47-68) 47/81	72 <sup>*4</sup> (63-81) 66/92	97 <sup>*4</sup> (90-100) 29/30	65 (52-76) 37/57	75 (64-85) 47/63	93 (81-100) 14/15
<b>LAM strip test (grade 2 cut-point) in smear-negative/ sputum scarce <i>M.tb</i> culture positive patients</b>	0 (0-26) 0/11	N/A	N/A	51 (36-66) 21/41	N/A	N/A	58 (41-74) 18/31	N/A	N/A
<b>Combined sputum smear microscopy and LAM strip test (grade 2 cut-point)</b>	58 (39-75) 15/26	79 (65-94) 23/29	N/R	75 <sup>#2 #3</sup> (65-83) 61/81	70 <sup>#5</sup> (60-79) 64/92	N/R	77 <sup>#6</sup> (65-86) 44/57	73 <sup>#8</sup> (62-84) 46/63	N/R

<sup>†</sup>14 of the 242 HIV-infected patients had no CD4 cell count data, and thus a total of 228 patients were available for this analysis. 107 *M.tb* culture positive group 1 patients were used for the sensitivity calculations and 121 *M.tb* culture negative group 1 and 84 non-TB group 2 patients were used for specificity calculations. The 120 patients in the CD4≤100 cells/ml group are a sub-group of those with CD4 ≤200 cells/ml.

<sup>‡</sup>Rapid (within 24 hours of hospitalization) smear microscopy includes both sputum and non-sputum samples (e.g. fine-needle aspirate and cerebrospinal fluid samples)

<sup>&</sup>2 CD4 >200 cells/ml and 2 CD4 ≤200 cells/ml group 1 patients were missing a LAM ELISA result.

\*Indicates  $p < 0.05$  for a comparison of differences in diagnostic accuracy measures between analyses and patient groups for a specific test (LAM ELISA or LAM strip test); specific p-values: <sup>#1</sup> $p < 0.001$ ; <sup>#2</sup> $p = 0.006$ ; <sup>#3</sup> $p = 0.002$ ; <sup>#4</sup> $p = 0.004$ ; <sup>#5</sup> $p = 0.01$ ; non-significant p-values not shown)

<sup>#</sup>Indicates  $p < 0.05$  for a comparison of the specific measures of diagnostic accuracy (sensitivity or specificity) between different tests (LAM ELISA vs. LAM strip test) or combinations thereof, or between different cut-points of the urine LAM strip test (grade-1 vs. grade-2 cut-point); specific p-value: <sup>#1</sup> $p = 0.02$ ; <sup>#2</sup> $p = 0.008$ ; <sup>#3</sup> $p = 0.02$ ; <sup>#4</sup> $p = 0.005$ ; <sup>#5</sup> $p < 0.001$ ; <sup>#6</sup> $p = 0.01$ ; <sup>#7</sup> $p = 0.003$ ; <sup>#8</sup> $p = 0.01$ ; non-significant p-values not shown)

N/R: no results; N/A: not applicable

University of Cape Town

Table B2. Performance outcomes of smear microscopy, LAM ELISA, and LAM strip test (using grade 1 and 2 cut-points) in HIV-infected patients stratified by CD4 count<sup>†</sup> (>200, ≤200 and ≤100 cells/ml) for analysis 2.

Diagnostic test(s)	HIV-infected patients with CD4 count >200 cells/ml			HIV-infected patients with CD4 count ≤200 cells/ml			HIV-infected patients with CD4 count ≤100 cells/ml		
	N=109			N=177			N=113		
	Sensitivity (95% CI)	Specificity <sup>§</sup> (95% CI)	Positive likelihood ratio (95% CI)	Sensitivity (95% CI)	Specificity <sup>§</sup> (95% CI)	Positive likelihood ratio (95% CI)	Sensitivity (95% CI)	Specificity <sup>§</sup> (95% CI)	Positive likelihood ratio (95% CI)
Rapid smear microscopy <sup>‡</sup>	38 (25-52) 17/45	98 (92-100) 63/64	24.2 (2.8-207.5)	44 <sup>#1</sup> (36-53) 58/131	100 (92-100) 46/46	N/C	48 <sup>#4</sup> (38-58) 41/86	100 (88-100) 0/27	N/C
LAM ELISA <sup>&amp;</sup>	28 <sup>*1</sup> (17-43) 12/43	100 (94-100) 63/63	N/C	59 <sup>*1</sup> (50-67) 76/129	96 (85-99) 43/45	13.3 (4.9-36.0)	68 (57-77) 57/84	92 (76-98) 24/26	8.8 (3.3-23.9)
LAM strip test (grade 1 cut-point)	49 (35-63) 22/45	92 <sup>#3</sup> (83-97) 59/64	6.3 (3.9-10.2)	64 (56-72) 84/131	89 (77-95) 41/46	5.9 (3.9-8.8)	69 (58-77) 59/86	89 (72-96) 24/27	6.2 (3.2-12.1)
LAM strip test (grade 2 cut-point)	29 <sup>*2</sup> (18-43) 13/45	100 <sup>#3</sup> (94-100) 64/64	N/C	52 <sup>*2 #2</sup> (43-60) 68/131	96 (86-99) 44/46	11.9 (4.4-32.7)	58 <sup>#5</sup> (48-68) 50/86	93 (77-98) 25/27	7.9 (2.9-21.5)
LAM strip test (grade 2 cut-point) in smear-negative/ sputum scarce <i>M.tb</i> culture positive patients	14 (6-32)	N/A	N/A	43 (32-54) 31/73	N/A	N/A	49 (35-63) 22/45	N/A	N/A
Combined sputum smear microscopy and LAM strip test (grade 2 cut-point)	47 <sup>*3</sup> (33-61) 21/45	98 (92-100) 63/64	29.9 (3.8-235.9)	68 <sup>*3 #1 #2</sup> (60-75) 89/131	96 (89-99) 44/46	15.6 (5.8-42.1)	73 <sup>#4 #5</sup> (63-82) 63/86	93 (77-98) 25/27	9.9 (3.7-26.7)

<sup>†</sup> 16 of the 302 HIV-infected patients had no CD4 cell count data, and thus a total of 286 patients were available for this analysis. 176 definite- or probable-TB were used for sensitivity calculations and 110 non-TB (26 group 1 and 84 group 2) were used for specificity calculations. Unclassifiable-TB patients were excluded from this analysis. The 113 patients in the CD4≤100 cells/ml group were a sub-group of those with CD4 ≤200 cells/ml

<sup>§</sup> Tabulated LAM strip test specificities for all CD4 cell-count sub-groups were calculated using group 1 and 2 non-TB patients combined. In the CD4 ≥ 200 cells/ml sub-group the specificity for group 1 and 2 using the grade 1 cut-point were 100% (72-100) and 91% (80-96), respectively; using the grade 2 cut-point they were 100% (72-100) and 100% (93-100), respectively. In the CD4 ≤200 cells/ml sub-group the relevant values using the grade 1 cut-point were 94% (72-99) and 87% (70-95), respectively; using the grade 2 cut-point they were 94% (72-99) and 97% (83-99), respectively. The relevant values for the CD4 ≤ 100 cells/ml sub-group using the grade 1 cut-point were 92% (65-99) and 87% (62-96), respectively, and using the grade 2 cut-point were 92% (65-99) and 93% (70-99), respectively. No significant inter-group specificity differences (1 vs. 2 in each CD4 count sub-group) was noted. 228

<sup>‡</sup> Rapid smear microscopy includes both sputum and non-sputum samples (e.g. fine-needle aspirate and cerebrospinal fluid samples) attainable within 24 hours of hospital admission.

&3 patients with a CD4 >200 (2 definite- or probable-TB and 1 non-TB) and 3 patients with a CD4 ≤200 (2 definite- or probable-TB and 1 non-TB) had a missing LAM ELISA result and 1 patient with a CD4 >200 (1 definite- or probable-TB) had a missing LAM strip test result.

\*Indicates p<0.05 for a comparison of differences in diagnostic accuracy measures between analyses and patient groups for a specific test (LAM ELISA or LAM strip test); specific p-values: <sup>#1</sup>p <0.001; <sup>#2</sup>p =0.008; <sup>#3</sup>p =0.01; non-significant p-values not shown)

#Indicates p<0.05 for a comparison of the specific measures of diagnostic accuracy (sensitivity or specificity) between different tests (LAM ELISA vs. LAM strip test) or combinations thereof, or between different cut-points of the urine LAM strip test (grade-1 vs. grade-2 cut-point); specific p-value: <sup>#1</sup>p<0.001; <sup>#2</sup>p=0.008; <sup>#3</sup>p=0.02; <sup>#4</sup>p<0.001; <sup>#5</sup>p=0.04; non-significant p-values not shown)

N/A: not applicable; N/C: unable to be calculated

University of Cape Town

*Inter-observer agreement using LAM strip grade 1 and grade 2 cut-points*

1.1 Reader 1 and 2 for the HIV-infected cohort patients (group 1 and 2 patients combined)  
[368 LAM strip tests]

		Reader 1		
		LAM Grade	0	1
Reader 2	0	213	12	1
	1	24	18	1
	$\geq 2$	3	13	83

Grade 1 cut-point Kappa = 0.77

Grade 2 cut-point Kappa = 0.87

1.2 Statistical comparison of agreement between two readers at cut-point grade 1 and 2.

Grade 1 cut-point: agreement: 328/368; disagreement: 40/368

Grade 2 cut-point: agreement: 350/368; disagreement: 18/368

McNemar chi square test for difference between agreement at grade-1 and grade-2 cut-points (p=0.002)

2. Additional evaluation of inter-observer and intra-observer agreement using 4 independent readers (211 HIV-infected group 1 and 2 patients; definite-, probable- and non-TB patients only with unclassified patients excluded)

*Inter-observer agreement*

i) Reader 1 vs 2

		Reader 1		
		LAM Grade	0	1
Reader 2	0	98	24	0
	1	32	12	1
	$\geq 2$	0	7	37

Grade 1 cut-point Kappa = 0.45

Grade 2 cut-point Kappa = 0.88

ii) Reader 1 vs 3

		Reader 1		
		LAM Grade	0	1
Reader 3	0	106	24	0
	1	22	5	0
	$\geq 2$	2	14	38

Grade 1 cut-point Kappa = 0.52

Grade 2 cut-point Kappa = 0.78

iii) Reader 1 vs 4

		<b>Reader 1</b>			
		<b>LAM</b>	<b>0</b>	<b>1</b>	<b>≥2</b>
<b>Reader 4</b>	<b>Grade</b>				
	<b>0</b>	93	7	0	
	<b>1</b>	36	27	0	
	<b>≥2</b>	1	9	38	

Grade 1 cut-point Kappa = 0.59

Grade 2 cut-point Kappa = 0.85

iv) Reader 2 vs 3

		<b>Reader 2</b>			
		<b>LAM</b>	<b>0</b>	<b>1</b>	<b>≥2</b>
<b>Reader 3</b>	<b>Grade</b>				
	<b>0</b>	102	28	0	
	<b>1</b>	17	10	0	
	<b>≥2</b>	3	7	44	

Grade 1 cut-point Kappa = 0.53

Grade 2 cut-point Kappa = 0.87

v) Reader 2 vs 4

		<b>Reader 2</b>			
		<b>LAM</b>	<b>0</b>	<b>1</b>	<b>≥2</b>
<b>Reader 4</b>	<b>Grade</b>				
	<b>0</b>	84	16	0	
	<b>1</b>	38	24	1	
	<b>≥2</b>	0	5	43	

Grade 1 cut-point Kappa = 0.49

Grade 2 cut-point Kappa = 0.92

vi) Reader 3 vs 4

		Reader 3		
LAM		0	1	$\geq 2$
Reader 4	Grade			
	0	88	10	2
	1	42	16	5
	$\geq 2$	0	1	47

Grade 1 cut-point Kappa = 0.49

Grade 2 cut-point Kappa = 0.90

*Intra-reader agreement*

(each reader grading the LAM strip test at 25 mins and again at 35 mins)

1. Reader 1

Kappa grade 1 cut-point: 0.67

Kappa grade 2 cut-point: 0.92

2. Reader 2

Kappa grade 1 cut-point: 0.69

Kappa grade 2 cut-point: 0.96

3. Reader 3

Kappa grade 1 cut-point: 0.73

Kappa grade 2 cut-point: 0.94

4. Reader 4

Kappa grade 1 cut-point: 0.56

Kappa grade 2 cut-point: 0.94

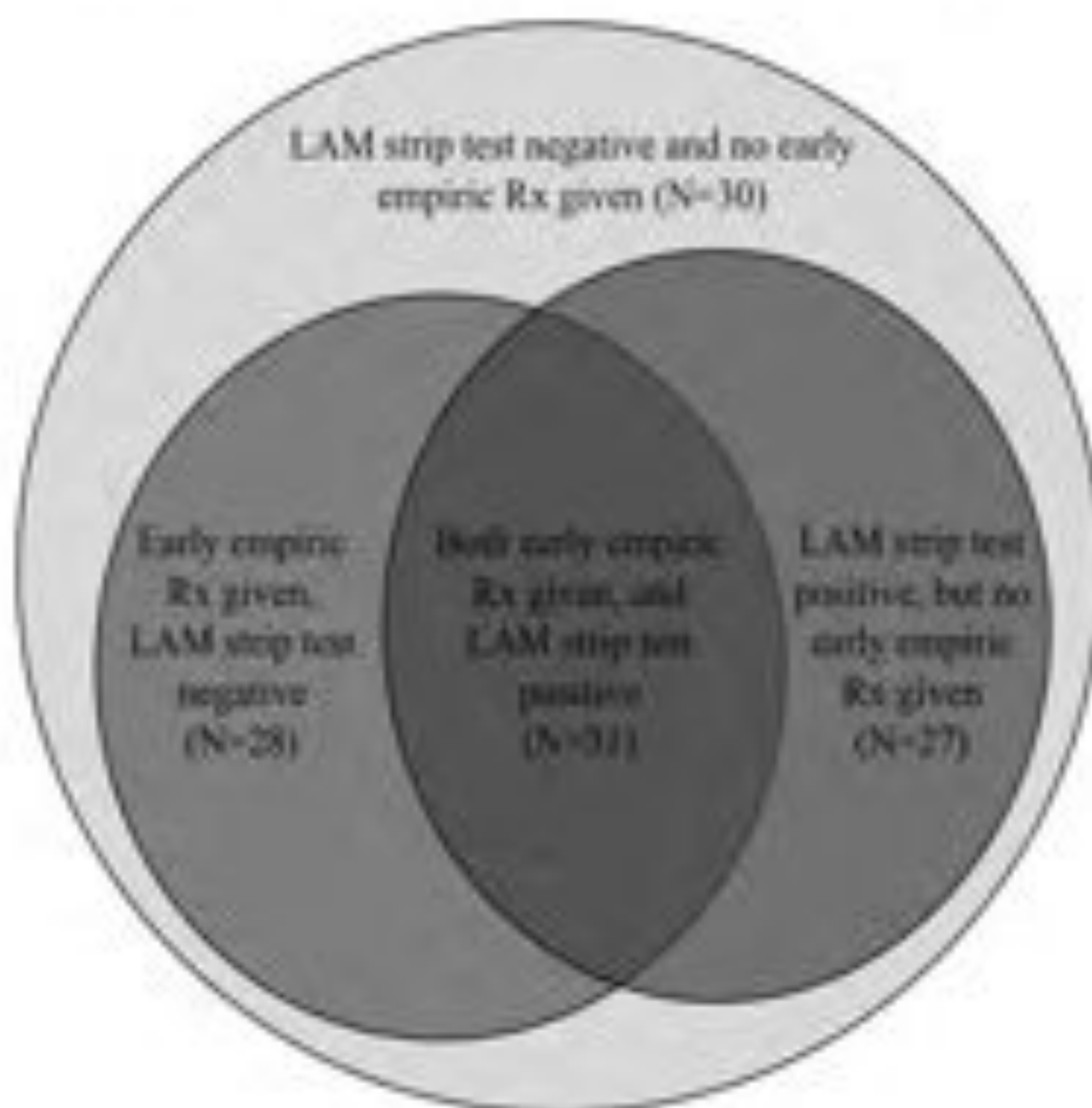
## 5. Chapter 9 supplementary material

Peter, J.G., Theron, G. and Dheda, K. "Can point-of-care urine LAM strip testing for tuberculosis add value to clinical decision making in hospitalised HIV-infected persons?"

2012 *PLoS ONE* 8(2): e54875

University of Cape Town

A. All definite TB (*M. tuberculosis* culture positive) (N=116)



Patient characteristic(s)	LAM strip negative but early empiric Rx given (N=28) Median (IQR) or n(%)	LAM strip positive, but no early empiric Rx given (N=27) Median (IQR) or n(%)	LAM strip negative and no early empiric treatment given (N=30) Median (IQR) or n (%)	P-value
Age, years	33 (24-38)	31 (26-38)	35 (31-40)	n/s
Previous TB	9 (32)	8 (30)	8 (27)	n/s
Weight, kgs	50 (44-60)	49 (45-56)	53 (46-62)	n/s
Temperature, °C	37.0 (36.5-38.0)	37.0 (36.0-37.5)	37.0 (36.5-37.8)	n/s
Respiratory rate, breaths/min	22 (19-29)	24 (20-28)	24 (20-28)	n/s
CD4 cell count, cells/ml	107 (54-171) <sup>*</sup>	62 (31-110) <sup>***1</sup>	180 (74-308) <sup>*1</sup>	<sup>*</sup> 0.05 <sup>**</sup> 0.002
MEWS	3 (1-4) <sup>*</sup>	5 (3-7) <sup>*</sup>	4 (3-6)	<sup>*</sup> 0.01
Urea, mmol/l	3.7 (3.4-4.6) <sup>*</sup>	5.5 (3.7-11.2) <sup>*</sup>	3.8 (2.7-7.6)	<sup>*</sup> 0.02
Creatinine, µmol/l	65 (52-80)	89 (50-105)	69 (52-100)	n/s
8-week mortality	3/24 (13)	6/25 (24) <sup>*</sup>	1/23 (4) <sup>*</sup>	<sup>*</sup> 0.05

**Table S1.** Diagnostic accuracy measures of early empiric treatment, the urine LAM strip test and CXR for TB diagnosis in hospitalised HIV-infected patients using *M. tuberculosis* culture positive-TB patients for sensitivity, and culture negative patients for specificity analyses.

<b>Diagnostic method</b>	<b>Sensitivity (%) (95% CI)</b>	<b>Specificity (%) (95% CI)</b>	<b>PPV (%) (95% CI)</b>	<b>NPV (%) (95% CI)</b>	<b>LR+ (95% CI)</b>
<b>Early empiric Rx<sup>†</sup></b>	51 <sup>*1*2</sup> (42-60) 59/116	63 <sup>*5*6*7</sup> (54-71) 79/126	56 <sup>*8</sup> (46-65) 59/106	58 (50-66) 79/136	1.36 (1.27-1.47)
<b>Urine LAM grade 2 cut-point)</b>	50 <sup>*3*4</sup> (41-59) 58/116	75 <sup>*5</sup> (67-82) 95/126	65 (55-74) 58/89	63 (57-69) 95/153	2.03 (1.84-2.24)
<b>CXR</b>	92 <sup>*1*3</sup> (86-96) 107/116	21 <sup>*7</sup> (15-29) 27/126	52 (45-59) 107/206	75 (58-86) 27/36	1.17 (1.15-1.20)
<b>Early empiric Rx plus urine LAM (grade 2 cut-point)</b>	74 <sup>*2*4</sup> (66-81) 86/116	75 <sup>*6</sup> (67-82) 95/126	74 <sup>*8</sup> (65-81) 86/117	76 (68-83) 95/125	3.01 (2.81-3.24)

<sup>§</sup> All patients with 1 or more valid *M. tuberculosis* culture (either sputum or non-sputum) are included in this analysis irrespective of final TB diagnostic categorization (39/281 patients excluded with either no/contaminated culture result)

<sup>†</sup> Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group.

P-values indicate significant differences between tests (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures

\*1 p<0.001; \*2 p<0.001; \*3 p<0.001; \*4 p<0.001; \*5 p=0.03; \*6 p=0.03; \*7 p<0.001; \*8 p=0.005

**Table S2.** Diagnostic accuracy measures of early empiric treatment, the urine LAM strip test and CXR for TB diagnosis in hospitalized HIV-infected patients using definite-TB (*M. tuberculosis* culture positive) for sensitivity, and non-TB patient groups for specificity analyses.

Diagnostic method	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	PPV (%) (95% CI)	NPV (%) (95% CI)	LR+ (95% CI)
Early empiric Rx <sup>†</sup>	51 <sup>*1*2</sup> (42-60) 59/116	100 (88-100) 27/27	100 (94-100) 59/59	32 (23-43) 27/84	N/C
Urine LAM (grade 2 cut-point)	50 <sup>*3*4</sup> (41-59) 58/116	96 <sup>*5</sup> (82-99) 26/27	98 (91-100) 58/59	31 (22-42) 26/84	13.5 (1.8-99.1)
CXR	92 <sup>*1*3</sup> (86-96) 107/116	30 <sup>*5</sup> (16-49) 8/27	85 (78-90) 107/126	47 (26-69) 8/17	1.3 (1.2-1.5)
Early empiric Rx plus urine LAM (grade 2 cut-point)	74 <sup>*2*4</sup> (66-81) 86/116	96 (82-99) 26/27	99 (94-100) 86/87	46 (34-59) 26/56	20.0 (2.8-143.2)

<sup>†</sup>Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group.

P-values indicate significant differences between tests (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures

\*1 p<0.001; \*2 p<0.001; \*3 p<0.001; \*4 p<0.001; \*5 p<0.001

**Table S3.** Univariate and multivariate analyses for associates of definite-TB in HIV-infected hospitalised patients.

Patient characteristic	Univariate analysis			
	OR (95% CI)	P-value	$\beta$ -coefficient	Score
Previous TB	0.57 (0.33-0.97)	0.04	n/a	n/a
Weight	0.97 (0.95-0.99)	0.01	n/a	n/a
Weight $\leq 50$ kg <sup>†</sup>	2.1 (1.2-3.7)	0.008	n/a	n/a
Temperature	1.5 (1.1-2.0)	0.005	n/a	n/a
Temperature $\geq 37.5$ °C <sup>†</sup>	2.6 (1.4-4.7)	0.002	n/a	n/a
CXR potentially TB	3.0 (1.4-6.9)	0.007	n/a	n/a
LAM Ag rapid test (cutpoint1)	3.8 (2.2-6.4)	<0.001	n/a	n/a
Multivariate analysis (clinical and radiology predictors only)				
History of previous TB	0.52 (0.29-0.93)	0.03	-0.65	-0.5
Weight $\leq 50$ kg	2.5 (1.4-4.6)	0.002	0.92	1
Temperature $\geq 37.5$ °C	2.8 (1.5-5.2)	0.001	1.02	1
CXR compatible with TB	3.0 (1.3-7.0)	0.01	1.09	1

<sup>†</sup>Receiver operating characteristic (ROC) curve-selected cut-point maximizing discriminatory utility used to dichotomise the continuous variables weight and temperature

OR: odds ratio; TB: Tuberculosis; CXR: Chest x-ray; LAM: Lipoarabinomannan

**Table S4.** Diagnostic accuracy measures for set of clinical predictors (using three ROC-selected cut-points), the urine LAM strip test and routine early empiric treatment in hospitalised HIV-infected patients using the definite and probable-TB groups for sensitivity and the non-TB groups for specificity analyses.

Type of TB detection test	Cut-point value	Sensitivity (%) (95% CI)	Specificity (%) (95% CI)	LR+ (95% CI)
Urine LAM	Grade 2	46 <sup>*1</sup> (39-53) 85/187	96 (92-99) 26/27	12.3 (1.7-89.6)
Early empiric Rx <sup>†</sup>	n/a	50 <sup>*1*2</sup> (43-57) 93/187	100 (88-100) 27/27	N/C
<b>Quantified set of clinical predictors</b>				
'rule-out'	≥ 0.5	91 (87-95) 171/187	15 (6-33) 4/27	1.07 (0.99-1.17)
Youden's index <sup>†</sup>	≥ 1.5	57 (50-64) 106/187	67 (48-81) 18/27	1.7 (1.35-2.14)
'rule-in'	≥ 2.5	10 <sup>*1</sup> (7-15) 19/187	100 (88-100) 27/27	N/C

P-values indicate significant differences between tests and/or cut-points (marked with \* and number to indicate comparison group) for different diagnostic accuracy measures; <sup>\*1</sup>p<0.001; <sup>\*2</sup>p=0.03<sup>†</sup> Youden's index is defined as the point on the ROC curve that provides the optimal mathematical balance between sensitivity and specificity

**Table S5.** Demographic, clinical, sampling and microbiological characteristics of study patients stratified by TB diagnostic group

	All N=281	Early empiric Rx given <sup>†</sup> N=120	No early empiric Rx N=161	P-value
<b>Demographics</b>				
Median age (yrs, IQR)	35 (29-39)	35 (30-40)	34 (29-39)	n/s
Female (n, %)	174 (62)	69 (58)	105 (65)	n/s
Median CD4 count (cells/ml)	86 (46-198)	84 (47-171)	93 (45-212)	n/s
Previous TB (n, %)	97 (35)	42 (35)	55 (34)	n/s
Current Smoker (n, %)	52 (19)	26 (22)	26 (16)	n/s
<b>Clinical features</b>				
Cough >2wks (n, %)	231 (82)	108 (90)*	123 (76)*	*0.003
Night sweats (n, %)	185 (66)	77 (64)	108 (67)	n/s
Self-reported Weight loss (n, %)	247 (88)	104 (87)	143 (89)	n/s
Fever >38°C (n, %)	49 (18)	22 (19)	27 (17)	n/s
Median weight (kgs, IQR)	53 (47-63)	52 (45-62)	53 (48-63)	n/s
Median temperature (°C, IQR)	36.8 (36.1-37.5)	36.9 (36.2-37.5)	36.8 (36.1-37.5)	n/s
Respiratory rate (breaths/min, IQR)	22 (19-28)	23 (20-28)	22 (19-28)	n/s
Median MEWS <sup>§</sup> (IQR)	4 (2-5)	3 (1-5)*	4 (3-5)*	0.001
Median urea (mmol/l, IQR)	4.8 (3.5-8)	4.5 (3.6-7.5)	5.3 (3.5-8.9)	n/s
Median creatinine (µmol/l, IQR)	72.5 (57-100)	68 (56-94)*	77 (59-107)*	0.04
CXR compatible with TB (n, %)	215 (77)	98 (82)	117 (73)	n/s
LAM strip test positive (grade 2) (n, %)	98 (35)	50 (42)*	48 (30)*	*0.04
<b>Clinical samples collected for TB culture</b>				
1 sputum sample (n, %)	207 (74)	91 (76)	116 (72)	n/s
≥2 sputum samples (n, %)	92 (33)	35 (29)	57 (35)	n/s
1 non-sputum sample (n, %)	160 (57)	67 (56)	93 (58)	n/s
≥2 non-sputum sample (n, %)	56 (20)	24 (20)	32 (20)	n/s
No samples (n, %)	19 (7)	5 (4)*	14 (8)	n/s

† Any patient commenced on TB treatment within 24 hours of hospital admission based only on clinical and radiological findings, and prior to the availability of any smear or culture results, is included in this group. Analysis is performed for all patients in this graph and hence includes 27 unclassified patients whom were commenced on early empiric treatment but do not form part of the primary analysis presented in the main manuscript.

P-values indicate significant differences between patient groups (marked with \* and number to indicate comparison group) for different patient characteristics

§MEWS: Modified early warning score is an admission triage score based on illness severity and higher scores correlated with poor outcomes and increased mortality [19]

University of Cape Town