Improving tuberculosis diagnosis in vulnerable populations: impact and cost-effectiveness of novel, rapid molecular assays

Hojoon Sohn

Department of Epidemiology, Biostatistics, and Occupational Health
Faculty of Medicine
McGill University, Montreal

February 2016

A thesis to be submitted to McGill University in partial fulfillment of
the requirements of the degree of Doctor of Philosophy (PhD)

Montreal, Quebec, Canada

TABLE OF CONTENTS

Front Matter

Abstract

Tuberculosis (TB) remains a major global public health concern with more than 9.6 million TB cases and 1.2 million deaths reported in 2014. The reduction of time delays in diagnosis and treatment of TB are the most important priorities in the control of the global TB epidemic, but shortfalls in conventional diagnostics of TB have significantly compromised this fundamental objective. Recently, the development of novel, rapid nucleic acid amplification tests (NAATs)[1] has generated great anticipation as a new strategy in addressing this important issue in TB control; however, the impact of NAATs on patient care improvement has not yet been well-characterized. With an ambitious End TB Strategy that aims to reduce TB incidence by 90% by 2035 and provide universal drug-susceptibility testing (DST), a comprehensive understanding of the evidence on test accuracy, impact of time delays in TB patient care, and cost and cost-effectiveness of these novel TB diagnostics are required to better guide relevant policy decisions.

The overall objective of this manuscript-based PhD thesis, therefore, was focused on a comprehensive assessment of the impact (diagnostic accuracy, patient and clinically important impact, cost and cost-effectiveness) of novel rapid molecular diagnostic tests for TB that have promising potential to improve TB patient care. This thesis is presented based on three manuscripts:

1. Impact of molecular diagnostic tests on TB diagnostic and treatment delays: a systematic review

In this first manuscript, we conducted a systematic review and meta-analysis to critically summarize and assess the impact of NAATs approved by the World Health Organization (WHO) on diagnostic and therapeutic delays compared to the standard diagnostic methods for drug susceptible (DS) and drug-resistant (DR) TB. We used both narrative and quantitative methods. We first developed a conceptual framework for defining time delays from the onset of disease symptoms to the initiation of anti-TB treatment. We also developed a checklist to assess and validate the quality of time delays reported based on five important methodologic

[1] Xpert MTB/RIF (Xpert), Line Probe Assay (LPA), or Loop-mediated isothermal amplification system for TB (TB-LAMP)

and contextual components to determine usefulness and comparability of the time delays. A total of 39 eligible studies evaluating time delay impact of NAATs (Xpert and LPA) were included in our review. Despite general evidence that the molecular TB diagnostics improved time delays in TB care, we found significant methodological inconsistencies and inadequacies in defining, evaluating, and reporting time delays across all studies. Subsequently, in this review, we elaborated on the importance of societal and health system factors associated with the time delays in TB care. Finally, we discussed how future research on time delays can utilize improved methodologies and reporting to properly inform future policy and implementation strategies of these next generation assays for TB diagnosis.

2. Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact

In this second manuscript, we studied diagnostic accuracy, feasibility, and potential patient-relevant impact of a point-of-care (POC) Xpert® MTB/RIF (Xpert), the first automated molecular test for TB, in a low-TB-burden, high resource setting. Our study of 502 patients enrolled in a university hospital TB clinic in Montreal, Canada showed that Xpert had a diagnostic accuracy much lower than those reported in the settings with high disease burden. As Xpert was not approved for clinical use in Canada, we assessed its hypothetical potential impact on time delays in diagnosis and treatment of TB and found that POC-Xpert could reduce delays in laboratory-based diagnosis of TB, but would have minimal impact on treatment delays. We concluded that the role of POC-Xpert was limited in the context of less extensive disease and in settings where the health system is well-resourced, and operational, diagnostic, and clinical guidelines are well-established.

3. Cost, cost-effectiveness, and affordability of the loop-mediated isothermal amplification (LAMP) assay for TB in high TB burden, low resource settings

In this third manuscript, the main objective was to assess cost-effectiveness of TB-LAMP, a novel molecular diagnostic test, in high TB burden, low resource settings. We used decision analytic modelling to simulate a cohort of patients diagnosed with TB in settings where Xpert MTB/RIF (Xpert) testing is not offered, parameterized for two countries with lower multi-drug resistant tuberculosis (MDR-TB prevalence): Malawi and Vietnam. We compared two TB-LAMP strategies, replacement for sputum smear microscopy (SSM) and add-on test to

SSM in smear negative patients, to the base case algorithm with SSM followed by clinical diagnosis in those patients with negative SSM. We performed a wide range of sensitivity analyses (one or two way, probabilistic, and scenario sensitivity analysis) to test the robustness of our findings. Our results indicate that both TB-LAMP strategies were cost-effective when comparing to the World Health Organization (WHO) willingness-to-pay (WTP) threshold levels. These conclusions did not change in the range of sensitivity analysis performed in our study. However, given TB-LAMP's lack of capacity to detect DR-TB, financial constraints in low income countries and emergence of novel automated point-of-care molecular tests for TB, policy makers must cautiously evaluate operational and financial feasibility prior to introducing this technology.

The future portfolio for TB diagnostics offers a great promise for much improved global TB control strategies; however, limited resources mandate that we evaluate them in both clinically and economically meaningful ways prior to their implementation in routine practice. This PhD thesis highlights the importance of a comprehensive assessment of diagnostic accuracy, patient relevant impact, cost and cost-effectiveness in evaluating novel diagnostic tests for TB, and provides relevant methods and empirical evidence that can better inform current and future directions of the decision making process in global TB control.

Résumé

La tuberculose (TB) demeure un problème de santé publique mondiale important, avec plus de 9,6 millions de cas et 1,2 millions de décès rapportés en 2014. La réduction des délais diagnostiques et de traitement de la TB représentent les priorités les plus importantes pour le contrôle global de l'épidémie, mais les lacunes des outils diagnostiques classiques ont compromis de façon significative ces objectifs fondamentaux. Les développements récents de tests rapides d'amplification des acides nucléiques (TAAN)[2] novateurs ont généré beaucoup d'enthousiasme en tant que nouvelles stratégies pour faire face au problème du contrôle de la TB. Toutefois, l'impact des TAAN sur l'amélioration des soins aux patients n'a pas encore été très bien caractérisé. Considérant l'ambitieuse stratégie mondiale de lutte contre la tuberculose (« End TB Strategy ») qui vise à réduire l'incidence de la TB de 90% avant 2035 et offrir un test universel de sensibilité médicamenteuse, une meilleure compréhension des

[2] Xpert MTB/RIF (Xpert), sondes en ligne ("Line Probe Assay" [LPA]), ou technique LAMP pour la TB ("Loop-mediated isothermal amplification system for TB" [TB-LAMP])

données sur l'exactitude diagnostique, l'impact des délais diagnostiques sur les soins aux patients, ainsi que le coût et le rapport coût-efficacité de ces nouveaux outils est nécessaire pour éclairer les décisions stratégiques liées à la TB.

L'objectif principal de cette thèse de doctorat par articles était donc l'étude détaillée de l'impact (exactitude diagnostique, paramètres d'importance pour les patients et le processus clinique, coût et rapport coût-efficacité) de tests diagnostiques moléculaires rapides pour la TB ayant le potentiel prometteur d'améliorer les soins pour les patients affectés par la TB. Cette thèse est basée sur trois articles:

1. Impact de tests diagnostiques moléculaires sur le diagnostic et le traitement de la TB: revue systématique ("Impact of molecular diagnostic tests on TB diagnostic and treatment delays: a systematic review")

Dans ce premier article, nous avons fait une revue systématique et méta-analyse afin de résumer et évaluer l'impact des TAAN approuvés par l'Organisation mondiale de la santé de façon critique, sur le délais diagnostiques et thérapeutiques comparés aux méthodes diagnostiques classiques pour la TB pharmacosensible ou résistante au médicaments. Nous avons utilisé des approches narratives et quantitatives. Nous avons d'abord développé un cadre conceptuel pour définir les délais à partir de l'apparition des symptômes de la maladie jusqu'à l'initiation du traitement anti-TB. Nous avons aussi développé une liste de vérifications pour évaluer et valider la qualité des délais rapportés, en se basant sur cinq composantes méthodologiques et contextuelles importantes afin de déterminer l'utilité et la comparabilité des délais. Un total de 39 études éligibles évaluant l'impact des délais avec des TAAN (Xpert et LPA) ont été inclus dans notre revue systématique. Malgré les résultats généralisés démontrant que les outils diagnostiques moléculaires pour la TB amélioraient les délais de soins, nous avons trouvé des incohérences et inadéquations méthodologiques significatives dans la définition, l'évaluation et les rapports de délais à travers toutes les études. Ainsi, dans notre revue systématique, nous avons élaboré sur l'importance des facteurs sociétaux et de santé associés avec les délais de soins pour la TB. Finalement, nous avons discuté comment la recherche future sur les délais pourrait utiliser des méthodologies et méthodes de rapport de données améliorées afin de mieux informer les politiques et stratégies de mise en application futures des prochaines générations de tests pour le diagnostic de la TB.

2. Utilisation de Xpert MTB/RIF dans un milieu à ressources élevées et faible incidence de TB: limitations en exactitude diagnostique et impact clinique ("Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact")

Dans ce second article, nous avons étudié l'exactitude diagnostique, la faisabilité et l'impact potentiel pour les patients de l'utilisation hors laboratoire de Xpert® MTB/RIF (Xpert), le premier test moléculaire automatisé pour la TB, dans un milieu à faible incidence de TB et ressources élevées. Notre étude de 502 patients recrutés dans une clinique de TB d'un hôpital universitaire de Montréal, Canada, a démontré que Xpert avait une exactitude diagnostique beaucoup plus faible que celle rapportée dans les milieux à plus forte incidence. Étant donné que Xpert n'était pas approuvé pour une utilisation clinique au Canada, nous avons évalué de façon hypothétique son impact potentiel sur les délais de diagnostic et de traitement de la TB, et avons trouvé que l'utilisation hors laboratoire de Xpert pouvait réduire les délais diagnostiques par rapport à une utilisation en laboratoire, mais que son impact sur les délais de traitement serait minimal. Nous avons ainsi conclu que le rôle de Xpert utilisé hors laboratoire était limité à des contextes où la maladie est moins présente, et où les directives opérationnelles, diagnostiques et cliniques du système de santé étaient bien établies.

3. Coût, rapport coût-efficacité, et abordabilité de la technique LAMP pour le diagnostic de la TB dans des milieux à forte incidence de TB et faibles ressources ("Cost, cost-effectiveness, and affordability of the loop-mediated isothermal amplification (LAMP) assay for TB in high TB burden, low resource settings")

Dans ce troisième article, l'objectif principal était d'évaluer le rapport coût-efficacité d'un nouveau test diagnostique moléculaire pour la TB, TB-LAMP, dans des milieux à forte incidence de TB et faibles ressources. Nous avons utilisé l'analyse décisionnelle modélisée pour simuler une cohorte de patients diagnostiqués pour la TB, dans des milieux où Xpert MTB/RIF (Xpert) n'était pas utilisé, paramétrés pour deux pays avec une faible incidence de TB résistante à de multiples médicaments: Malawi et Viêt Nam. Nous avons comparé deux stratégies TB-LAMP, un remplacement de l'examen microscopique des frottis de crachat et test additionnel à ce dernier dans le cas des patients à frottis négatif, à l'algorithme du cas de base, soit l'examen microscopique des frottis de crachat suivi d'un diagnostic clinique chez les patients dont le frottis d'avère négatif. Nous avons réalisé un éventail d'analyses de

sensibilité (à un ou deux sens, probabiliste, et scénario d'analyse de sensibilité) afin de vérifier la robustesse de nos résultats. Nos résultats démontrent que les deux stratégies TB-LAMP étaient rentables par rapport aux niveaux seuils de disposition à payer de l'Organisation mondiale de la santé. Ces conclusions sont demeurées les mêmes pour l'ensemble des analyses de sensibilité exécutées dans notre étude. Toutefois, étant donné l'incapacité des TB-LAMP à détecter la TB résistante aux médicaments, les limitations financières de pays à faible revenu, et l'émergence de nouveaux tests moléculaires automatisés pour la TB pouvant être utilisés hors laboratoire, les décideurs devront évaluer la faisabilité opérationnelle et financière avec prudence, avant d'introduire cette technologie.

Le portfolio futur des outils diagnostiques pour la TB offre un potentiel prometteur pour l'amélioration des stratégies de contrôle global de la TB. Toutefois, les ressources limitées impliquent qu'ils doivent être adéquatement évalués de façon clinique et économique avant de les mettre en application dans la pratique courante. Cette thèse de doctorat met en lumière l'importance d'une évaluation détaillée de l'exactitude diagnostique, les impacts importants pour les patients, le coût et le rapport coût-efficacité dans l'évaluation de nouveaux outils diagnostiques pour la TB, et fournit des méthodes appropriées ainsi que des preuves empiriques qui permettront de mieux éclairer la direction que devrait prendre le processus décisionnel présent et future dans le contrôle global de la TB.

Acknowledgements

I still remember vividly, one coincidental day in January of 2009 when I met Dr. Madhukar Pai in a remote rural city in India. What was meant to be a short exchange of greetings turned out to be a meeting that transformed my life completely. Underneath the calming natural shade protecting us from the strong South Indian sun, we discussed the future of TB research, how I can contribute to this rapidly changing field, and what it will take me to make this a reality. Our conversation that day culminated into my studies at McGill that were not only challenging, but also opened doors to countless opportunities for learning and career establishment. I am most grateful for Dr. Pai's endless moral and intellectual support throughout my days here at McGill. While I was struggling through many unforeseen turns of events in my personal and academic life, he was always there to help and guide me, and most importantly to tell me not to give up. Without you, I would not be the same person I am today.

I would also like to extend my gratitude to my supervisory committee members, Dr. Richard Menzies, Kevin Schwartzman, and Dr. Andrea Benedetti for their collaborative spirits and patience with me as I slowly paved my way through my studies at McGill. I am also grateful for all my international collaborators and fellow students for their hospitality, expertise, and most importantly making our partnerships a truly learning experience for all. Particularly, I would like to thank Samuel Schumacher for our intellectual conversations and working together on the systematic review. I would also like to thank Dr. Claudia Denkinger for helping me with the Xpert project in Montreal while I was serving in the Korean Army.

Lastly, I am most grateful to my wife, Mineui, for persevering through all the ups and downs during our life together, especially during the years when I was enlisted in the Korean Army in the middle of my doctoral research. Your endless love and support have truly been a motivation for me to overcome all the difficulties in reaching my goal. I also thank my parents for always believing in me and for their unconditional love.

For

My beautiful son, *Alex Haewon*

Fiat Mihi Secundum Verbum Tuum

Preface & Contribution of Authors

All of the co-authors of the manuscripts included in this thesis have read and agreed to their submission:

Manuscript 1: Impact of molecular tuberculosis diagnostics on reducing time delays in diagnosis and treatment of tuberculosis: a systematic review

H Sohn designed the study and developed objectives, in conjunction to M Pai and S Schumacher. **H Sohn** wrote the study protocol, designed data extraction forms, extracted data, performed all analyses, and drafted the manuscript. S Schumacher served as second reviewer of the included studies and independently extracted and verified the data. Z Qin served as second reviewer for screening eligible titles and abstracts, and G Gore helped develop search strategies and performed the search. A Benedetti provided input on meta-analysis and verified the analysis results. All co-authors gave H Sohn the permission to use this manuscript for this thesis.

Manuscript 2: Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact (published in *Clinical Infect Dis* 2014)

H Sohn wrote the CIHR grant that funded the project with M Pai. **H Sohn** and M Pai designed and developed the protocol for the study with input from D Menzies, M Behr, and K Schwartzman. **H Sohn** and C Denkinger implemented and coordinated the study and A Aero managed patient enrolment & data collection. C Denkinger performed analysis and **H Sohn** provided assistance with the data analysis. **H Sohn**, C Denkinger, and M Pai wrote the manuscript and all other authors critically reviewed the manuscript. All co-authors gave H Sohn the permission to use this manuscript for this thesis.

Manuscript 3: Cost-effectiveness of TB-LAMP assay for diagnosis of pulmonary tuberculosis

H Sohn was the principal investigator for this study and developed the protocol with inputs from A van't Hoog and F Cobelens. **H Sohn** designed the decision analytic model, collected all cost and effectiveness data, performed all coding and analysis, interpreted results, and drafted the manuscript. A van't Hoog and F Cobelens provided guidance on the development of the model, critically reviewed all of the coding in the model, reviewed the data used in the

model, assisted with the analysis, and critically reviewed the manuscript. **H Sohn** and F Cobelens participated in the development of the original model (1) (Vassall, 2011) that was adopted for this study. All co-authors gave H Sohn the permission to use this manuscript for this thesis.

Statement of Originality

The three manuscripts that are included in this thesis are of original scholarship and provide contributions to existing knowledge. In all 3 manuscripts, I am the first author. In the first manuscript (Chapter 3), we were the first group to perform a systematic review and meta-analysis on the impact of WHO-approved NAATs on reducing time delays for diagnosis and treatment of DR and DS-TB. In addition, we provided a revised framework to define and clarify key time delay components relevant to the diagnosis and treatment process of patients with symptoms of TB. We are also first to develop a criterion for evaluating the quality of reporting time delays. The second manuscript (Chapter 4) is one of the first studies to empirically address the impact of Xpert in low-incidence, high-resource settings with full mycobacterial culture and drug susceptibility testing capability. It is the first evaluation of Xpert MTB/RIF in Canada and one of the first studies to evaluate its impact at the point-of-care. This manuscript was published in *Clinical Infectious Disease* in 2014. The third manuscript (Chapter 5) is the first cost-effectiveness study for the TB-LAMP assay. This study was also the first to provide a scenario sensitivity analysis to assess cost-effectiveness of TB-LAMP used in conjunction with Xpert. The findings from this study were presented at the WHO Expert Group meeting on TB-LAMP assays and will be used for the WHO's policy guideline development for TB-LAMP.

Research Ethics

Ethics approval was not required for *Manuscript 1 and 3* (Chapter 3 and 5) as it was a systematic review and decision analysis. McGill University's Faculty of Medicine Institution Review Board Procedures approved the procedures of the study conducted for *Manuscript 2*. Participants were provided with written informed consent, which was recorded using a consent form reviewed and approved by McGill University's Faculty of Medicine Institutional Review Board. The signed research ethics statement from McGill University's Faculty of Medicine Institution Review Board are retained by the supervisor (Dr. M. Pai) in accordance with McGill's policies on research ethics.

List of Abbreviations

CEAC: Cost-effectiveness acceptability curve

CI: confidence interval

CXR: Chest X-ray or Chest Radiography

DR-TB: Drug resistance tuberculosis

DS-TB: Drug susceptible tuberculosis

DST: Drug Susceptibility Testing

GDP: Gross Domestic Product

GNI: Gross National Income

GRADE: Grading of Recommendations Assessment, Development and Evaluation

HIV: Human Immunodeficiency Virus

ICER: Incremental cost-effectiveness ratio

LAMP: Loop-mediated isothermal amplification

LED: Light-emitting diode

LMIC: Low and middle income countries

LPA: Line Probe Assay

MDR-TB: Multi-drug resistance tuberculosis

MTB: Mycobacterium tuberculosis

NAAT: Nucleic Acid Amplification Test

NTM: Non Tuberculous Mycobacteria

POC: Point-of-care

RCT: Randomized Controlled Trial

RIF: Rifampin

SSM: Sputum smear microscopy

TB: Tuberculosis

WHO: World Health Organization

WTP: Willingness to Pay

ZN: Zhiel Neelsen

List of Figure

Chapter 1: Introduction

1.1 Background

Tuberculosis (TB) is one of the world's most important infectious causes of morbidity and mortality among adults. An estimated one-third of the world's population is infected with TB, and nearly 9 million new TB cases and 1.5 million TB deaths occur each year (2). Although globally, TB has steadily declined each year since 2000, the majority of the disease burden continues to lie in low and middle income countries (LMICs) with the countries in the South-East Asian, Western Pacific, and African regions accounting for more than 80% of the total TB incidence reported in 2013 (2). Canada has one of the lowest overall TB incidences world-wide (3); however, its TB problem is unevenly spread across the country with TB rates astoundingly high amongst immigrants from endemic countries and Aboriginal populations. These two populations accounted for more than 70% of all reported cases in 2013 (3).

Early case detection and prompt treatment are the most critical steps of the global TB control strategy (4). However, several studies have shown that diagnostic delays are very common (5, 6), partly because TB diagnosis continues to rely on century-old tests such as sputum smear microscopy and chest radiography. Consequently, more than one third of all TB cases were undiagnosed or unreported (2), underscoring the urgent need for a diagnostic tool that is accurate, can be placed at near the patient care, and is simple to operate. Such a tool would have significant impact on TB control through an interruption of transmission and potentially improving treatment outcomes in those diagnosed earlier (7, 8).

Against this backdrop, new sources of funding and innovative product development partnerships were formed (9), which has stimulated greater efforts in developing novel diagnostic platforms (10). These phenomena have led to development of commercialized nucleic acid amplification tests (NAATs), such as Line Probe Assays (LPA) and Xpert MTB/RIF (Xpert) that offer major advantages of speed and sensitivity. WHO's endorsements of these technologies (11, 12) have subsequently led to unprecedented efforts to scale-up these technologies globally (13-15). However, these tests are expensive to perform and actual diagnostic coverage and utilization have been limited to intermediate or higher level laboratories, implying limited patient level impact, particularly on diagnostic and treatment delays.

With the growing number of innovative TB diagnostic tests on the horizon (10), it is important to address current knowledge gaps regarding NAAT accuracy and their greater effects; gaps of particular note are 1) test accuracy in low TB incidence settings, 2) feasibility and impact of innovative approaches in test implementation (e.g. point-of-care test setting) and 3) actual impact on patient care, particularly in reducing diagnostic and treatment delays. In addition, it is also critical to understand the cost-effectiveness of novel NAATs which could potentially bring incremental value in further improving the diagnostic capacity for TB in resource-limited settings, specifically in those areas where current NAAT platforms cannot be implemented. This knowledge will be vital in providing directions for policy and guideline development in the post-Xpert era.

1.2 Thesis aim and objectives

Therefore, the overarching goal of my thesis is to address these important knowledge 'gaps' to help shape future policy development of novel TB diagnostic tests. Throughout my thesis research, I focused on an array of issues and produced a wide range of scientific evidence, including evaluation of diagnostic accuracy, clinical and patient impact, and cost and cost-effectiveness of NAATs, which are essential to the policy decision making process. The aims and specific objectives of each chapter of my thesis research are as follows:

1. To systematically review the evidence regarding the reduction in diagnostic and therapeutic delays for drug-susceptible (DS) and drug-resistant (DR) forms of TB, with the use of the WHO-approved NAATs, and following specific objectives:
   A. To critically summarize and quantitatively assess the impact of the WHO-approved NAATs on time delay reduction for the diagnosis and treatment of DS and DR-TB
   B. To develop a conceptual framework for defining time delay components for TB diagnosis and treatment
   C. To propose a systematic approach for appraising methodological and reporting quality of time delay estimates

2. To study feasibility and diagnostic accuracy of the point-of-care Xpert in a low-incidence, high resource setting, with the following specific objectives:

A. To assess the technical performance (diagnostic accuracy, limits of detection) and feasibility of the point-of-care Xpert in a high throughput clinic specialized in TB with the test performed by healthcare workers who are not laboratory trained

B. To assess potential clinical impact of the point-of-care Xpert on reducing diagnostic and treatment delays

3. To assess cost-effectiveness of loop-mediated isothermal amplification assay for TB (TB-LAMP) implemented at peripheral laboratories in resource-poor, high TB-burden settings, with the following specific objectives:

A. To assess cost-effectiveness of TB-LAMP assay as a replacement or an add-on test to sputum smear microscopy (SSM) in settings with low multi-drug resistant TB where Xpert is not used as a routine diagnostic test, and in areas with low and high HIV prevalence

B. To assess cost-effectiveness of TB-LAMP assay as a replacement or an add-on test to SSM in settings with Xpert available as a routine diagnostic

Chapter 2: Literature review

2.1 Tuberculosis

2.1.1 Global epidemiology of tuberculosis

Tuberculosis (TB) is an airborne infectious disease caused by the bacteria *Mycobacterium tuberculosis* (MTB) that primarily infects the lungs (pulmonary TB), but can also infect other parts of the body (extra-pulmonary TB) (16). As one of the world's most important transmissible causes of morbidity and mortality among adults, TB accounts for an estimated one-third of the world's population, with nearly 9.6 million new TB cases and 1.2 million deaths occurring each year (17). Subsequently, someone in the world is newly infected with MTB every second and more than 2 billion people are infected with MTB in total.

Though the TB prevalence has been reduced by more than 42% from 1990 estimates (Figure 1a) (2), the TB disease burden continues to be concentrated in the 22 high endemic countries with no or limited signs of improvement in TB incidence in the vast majority of these countries (Figure 1b). Of these 22 high-burden countries, China, India, and Indonesia cumulatively accounted for more than 40% of the total global disease burden (Figure 2).

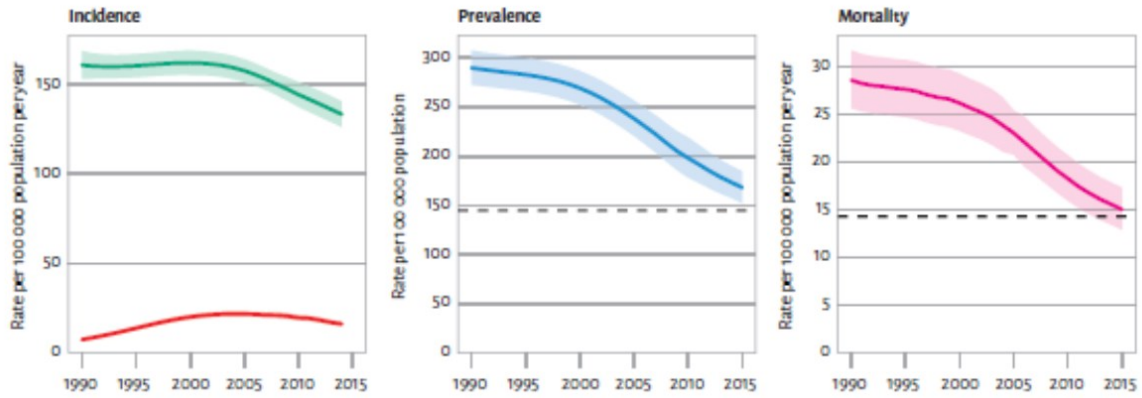


Figure 1a. Global trends in estimated rates of TB incidence (1990-2014), Global Tuberculosis Report 2015 (pg. 23).

* Left: estimated incidence rate include HIV-positive (green) and estimated incidence rate of HIV-positive (red). Center and right: The horizontal dashed lines represent the Stop TB Partnership targets of a 50% reduction in prevalence and mortality rates by 2015 compared with 1990. Shaded areas represent uncertainty bands. Mortality excludes TB deaths among HIV-positive people(2)

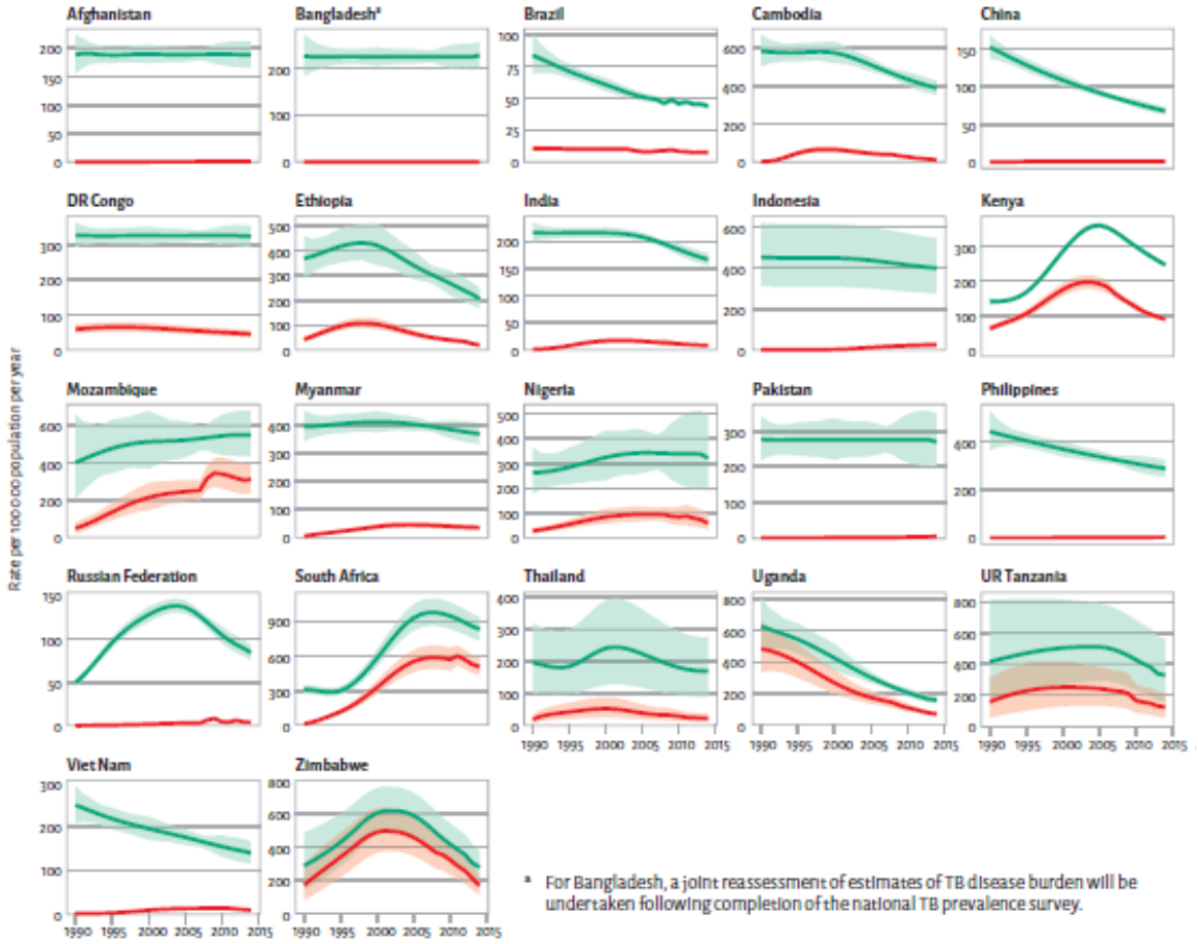


Figure 1b. Estimated TB incidence rates, 22 high-burden countries, 1990-2014, Global Tuberculosis Report 2015 (pg. 24).

* Estimated TB incidence rates (green) and estimated incidence rates of HIV-positive (red). Shaded areas represent uncertainty bands. (2)

In these countries, the problem of TB is compounded by poor management of active TB, which is a result of limited diagnostic tools and inconsistent supply and use of essential antibiotic drugs. Consequently, more than 37% of the TB and 75% of multi-drug resistant TB (MDR-TB), a rifampicin and isoniazid-resistant form of TB with or without resistance to other first-line drugs, were undiagnosed and unreported, potentially fueling a wide spread of antibiotic resistance strains, co-infections with HIV, and influencing other socio-economic factors, all of which result in great challenges for TB control programs in resource limited settings. Likewise, low case detection rates continue to be a major concern, and early

diagnosis and rapid treatment to interrupt transmission remain the top priorities for global TB control (18).

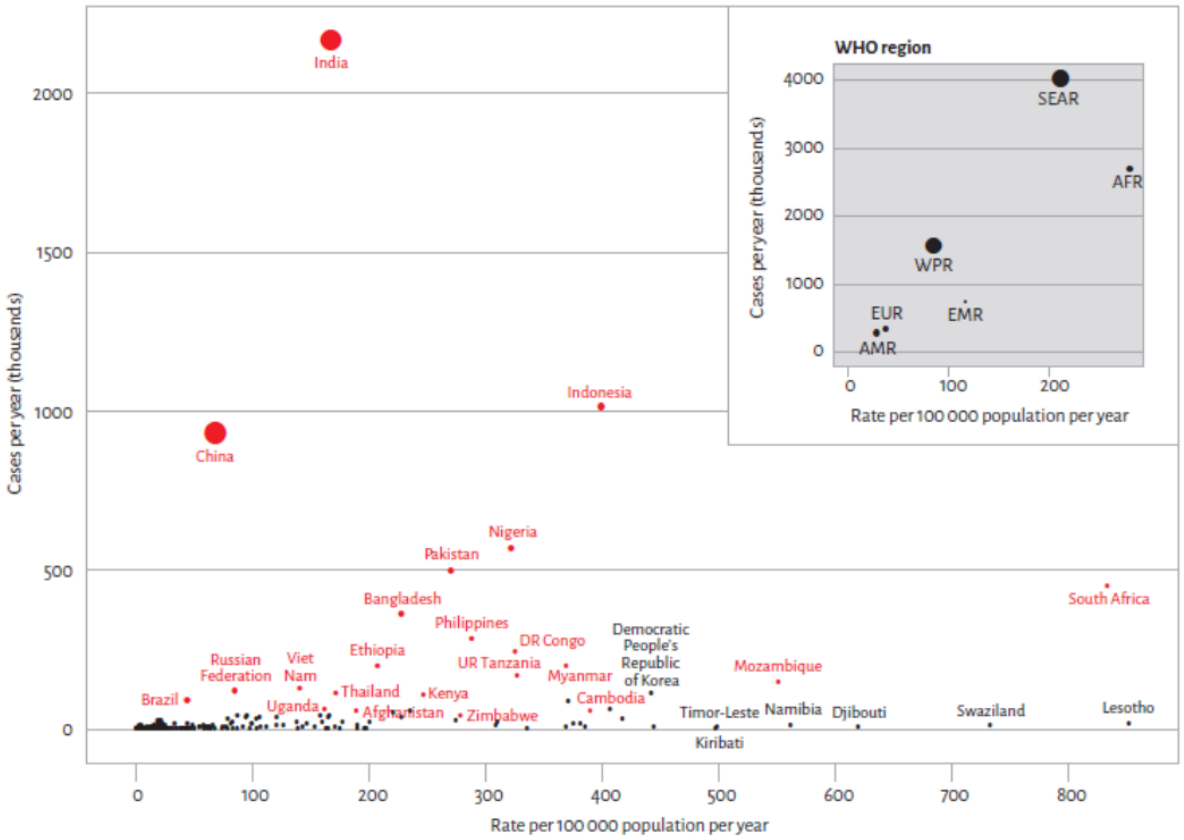


Figure 2. Global distribution of estimated TB incidence by rate and absolute number, 2014, Global Tuberculosis Report 2015 (pg. 19).

* The size of each bubble is proportional to the size of the country's population. High-burden countries are shown in red. (2)

2.1.2 Tuberculosis in Canada

TB continues to be an important health problem in Canada (19). While overall incidence continued to decrease to the low level of 4.7 per 100,000 persons in 2012 (Figure 3), this figure masks important disparities in rates (Figure 4) (19). Canada has seen consistent growth in the number of temporary and permanent residents entering and residing in Canada since 1985. With a considerable proportion of these immigrants coming from high TB endemic countries, the TB rate among foreign-born residents is more than 14.8 per 100,000, a rate 6 fold higher than the Canadian-born, non-Aboriginal population. In 2013, foreign-born individuals accounted for 71% of all reported TB cases in Canada (19). Drug-resistant TB is

very uncommon in Canada and is well below the global levels. In 2008, only 1.1% (15) of the 1359 isolates reported to the Public Health Agency of Canada (PHAC) were classified as multi-drug resistant (MDR-TB).

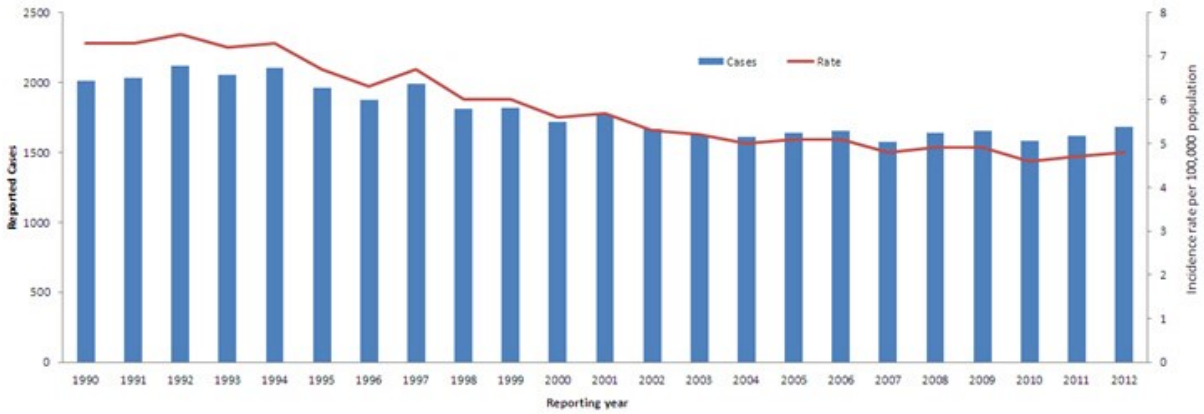


Figure 3. Annual number of reported tuberculosis cases and incidence rate in Canada 1990 – 2012 (3)

Figure 4. Annual number of reported tuberculosis cases and incidence rates by population: 2002 – 2012 (20)

TB incidence in Quebec was below the national level at 2.9 per 100,000 and new cases in Quebec accounted for approximately 15% of 1640 total reported cases in Canada in 2013 (3). More and more immigrants are choosing Quebec as their residence; in 2000, it was the third most settled province, and new immigrants in Quebec now account for more than 20% of all immigrants to Canada in 2013, making it the second most settled province, next to Ontario.

Within Quebec, Montreal accounted for more than 86% of immigrants settling into the province.

The TB rate in the Canadian-born Aboriginal population continues to be the highest of the three groups, at.19.9 per 100,000 accounting for 19% of the total cases reported in 2013 (19). Even compared to other Aboriginal groups in Canada, the Inuit population have a disproportionately high rate of TB, with the Inuit population in Nunavut having the highest TB incidence, at a rate of 154.2 per 100,000. The urgency of this problem is underlined by Nunavut having just recorded its highest incidence of TB in the past 10 years, with over 100 cases in 2010 (21).

2.2 Diagnosis of tuberculosis

2.2.1 Current challenges in diagnosis and care of tuberculosis

Early diagnosis and prompt treatment is the cornerstone of TB control (18). Although case detection is acknowledged to be the first critical step (4), several studies have shown that diagnostic delays are very common (5, 6) partly because TB diagnosis continues to rely on century-old tests, such as sputum smear microscopy and chest radiography. Thus, missed or delayed diagnosis results in ongoing transmission, mortality, and social and economic consequences. Lack of rapid, simple and accurate diagnostic tests is a major hurdle in controlling the global burden of TB. While this very problem is most significant in developing countries with high burden of TB, Canada, where there are ample resources and the best diagnostic tests available for TB, is not exempt from this problem, as existing TB diagnostic tests either are subject to their inherent limitations, or they cannot be performed at the point-of-treatment.

Smear microscopy: Direct (i.e. unconcentrated) sputum smear microscopy (where acid fast bacilli [AFB] are visualized using Ziehl-Neelsen [ZN] staining) remains the most widely used test to diagnose TB. Although the technique is inexpensive, specific and not technically demanding, it has several limitations including unacceptably low sensitivity in diagnosing TB and the inability to predict susceptibility to drugs used for treatment (22-24). Sputum smear microscopy (SSM) only detects TB cases with large numbers of mycobacteria in the sputum, explaining its poor sensitivity (~60%), which is particularly low for diagnosing TB associated with HIV infection as well as disease in children. In Canada, microscopy is done via the

auramine staining method using concentrated specimens, with the slides read by fluorescent microscopy (25). By definition, SSM has no role in the diagnosis of smear-negative TB cases, less contagious and the most common type of TB in low disease burden settings like Canada. To overcome the problem of low sensitivity, 3 sputum smears need to be performed for each patient, and this often results in diagnostic delays and patient drop outs (26). Even in highly organized clinical settings such as in Montreal, sputum smears take, on average, 2 - 4 days for results to be reported to the TB clinic.

Culture: Culture is the gold standard of diagnosis and is a highly sensitive method that allows identification of *MTB* as well as differentiation between drug-susceptible and resistant strains. However, its benefits are often outweighed by a long turn-around-time (TAT) of more than 2 - 4 weeks (6 - 8 weeks for solid culture) and the requirement of a bio-safety level 3 (BSL3) laboratory, which restricts the technology to reference level laboratories with negative air-pressure systems. While time to detection is only about 2 weeks with modern liquid-medium culture methods such as BACTEC MGIT 960 [BD Diagnostic Systems, NJ, USA], this test is expensive and not widely accessible. In Montreal, average turn-around-times for culture results is about 2 - 3 weeks. Likewise, considerable delays in diagnosis using culture considerably limits its utility for actual patient management, particularly in low-resource settings.

Conventional molecular tests: Nucleic acid amplification tests (NAAT) have the potential to detect *M. tuberculosis* DNA within hours, via amplification of *MTB*-specific nucleic acid sequences in clinical specimens (23, 27, 28). NAATs offer significantly increased speed compared to culture and are highly sensitive in sputum smear-positive samples (29). However, sensitivity tends to be low in smear-negative cases (29). The vast majority of conventional NAATs have complex, multi-step specimen processing and DNA extraction procedures, and require expensive, dedicated instruments with well-established laboratory infrastructure and highly trained laboratory staff (30), and as such, most of the conventional NAATs do not meet the needs of resource-poor countries. This severely limits their potential to impact patient management in high TB-burden settings.

2.2.2 The importance of reducing diagnostic and treatment delays

Delays in diagnosis and treatment are the most important impediments to effective TB control, and facilitates continued transmission of the infection, increases morbidity and

mortality of TB patients, and ultimately contributes to the slow and low decline in TB incidence (31). Several studies have shown that diagnostic delays are common, particularly in the low and middle income countries (LMICs) (6, 32, 33). This is a consequence of a reliance on ineffective diagnostics in addition to patient and health system related factors (5, 6). According to a systematic review by Sreeramareddy and coworkers, the median diagnostic delay days experienced by patients in LMICs was 28.4 days from the initial TB specific consultation , while total median health systems delay (initial symptom evaluation to TB specific treatment initiation) was 67.8 days (6). Similar findings were reported in a subsequent systematic review that focused on 23 studies from different parts in India, with a total health systems delay amounting to a median 55.3 days (34).

In a high income setting such as Montreal, diagnostic delay is also a problem, with many facets (35). Since most TB cases are sputum smear-negative, a positive culture is the only confirmatory test, and this can mean a delay of 2 - 3 weeks before correct treatment can be initiated. While smear-negative patients are not as infectious as smear-positive cases, they do contribute to disease transmission, accounting approximately for 17% of the overall TB transmission as reported in a study in Vancouver (36, 37). Another diagnostic challenge pertains to patients who are isolated because of suspected active TB. Currently, patients suspected to have active TB may sometimes be hospitalized in isolation rooms until cultures come back negative, and as this may take up to 2 - 3 weeks, considerable medical, social and financial burdens are placed on the patients, even those found to ultimately be TB negative (Figure 5). Similarly, unnecessary hospitalization worsens crowding, prolongs wait times, and poses an economic burden on hospitals and to the health system.

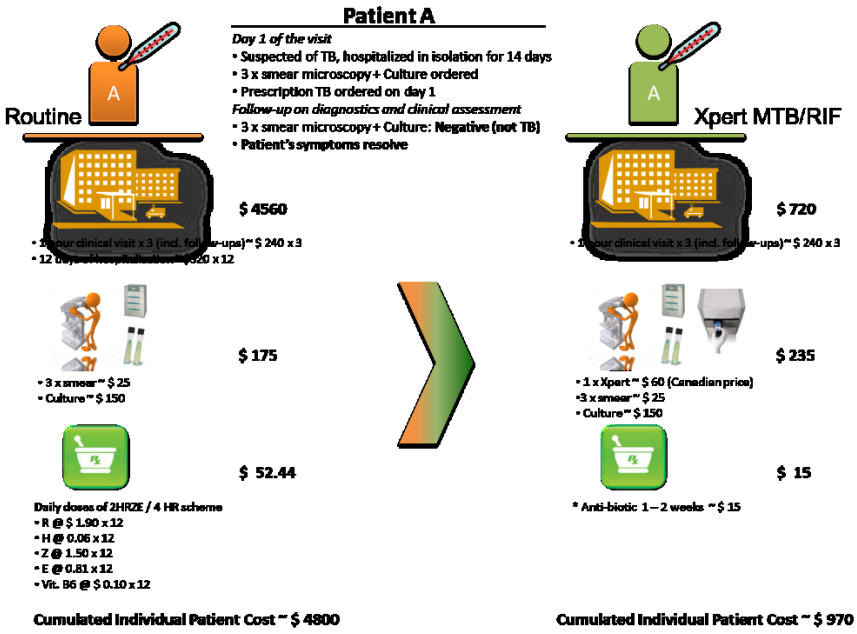


Figure 5. Illustrative example of cumulative individual patient costs comparing routine diagnosis and rapid point-of-care testing with Xpert.
* Costs are estimated based on the WHO-CHOICE cost estimates for hospital visits and hospitalization, and the costs for tuberculosis care in Canada.

In a modeling study examining factors that influence overall delays in TB diagnosis and treatment, it has been demonstrated that improved test accuracy, particularly in smear-negative TB patients, and access to care are key factors that can considerably improve time delays experienced by patients suspected of TB disease (38). Similarly, a systematic review assessing the factors influencing pre-treatment loss to follow-up in TB patients reported that health-system-related obstacles, particularly delays in receiving diagnostic results, were major reasons for these patients not starting TB treatment (39). Likewise, improving health system efficiency (e.g. same-day smear) or scaling up laboratory capacity to include novel diagnostic tests (e.g. rapid molecular diagnostic test) can potentially considerably impact reducing time delays and, ultimately, TB incidence (40).

2.2.3 Recent advancements in diagnostic tools for tuberculosis

Recently, rapid NAATs with technical designs suitable for use in LMICs have been developed and are being used in high TB-burden settings. The Xpert MTB/RIF® assay (Xpert) (Cepheid, Sunnyvale, CA, USA) is an automated NAAT with the capacity to detect both MTB and rifampicin resistance, a marker of multi-drug resistance (MDR) MTB strains, within two hours (Figure 6).

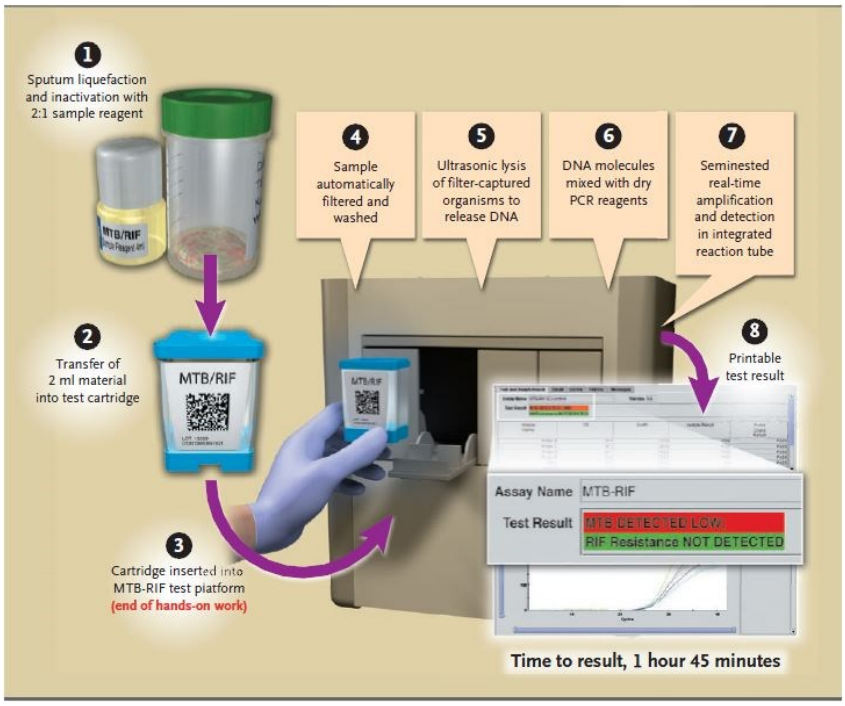


Figure 6. Xpert Testing process and sample preparation process for Xpert MTB/RIF, Boehme et al., *NEJM* 2010 (41)

Its test accuracy is equivalent to that of solid culture (89%, 95% CI: 85-92) and a specificity of 99% (95% CI: 98-99) (42) and the use of "pod" like disposable and closed configured cartridges eliminates the risk of cross-contamination and the risk of infectious-aerosol formation below that of smear preparation (43, 44). These characteristics of Xpert make it suitable for lower biosafety level laboratories and even for use in point-of-care (POC) settings, a true 'game changer' in TB diagnosis.

Since the WHO's endorsement of Xpert in 2010 (45), there has been a surge of donor funding and an establishment of a concessionary pricing scheme for LMICs (46, 47) to facilitate the global roll-out of Xpert, particularly in high TB-burden countries. As of December 2014, more than 3,700 GeneXpert instruments and 10 million Xpert MTB/RIF cartridges have been procured in the public sector (Figure 7) (48) . Despite these efforts, Xpert is still relatively costly to operate (49) and requires supportive infrastructure (e.g. steady electricity supply) and continued maintenance of the instrument, limiting its use at laboratories with limited infrastructure as a routine test. Likewise, there still are considerable gaps in Xpert diagnostic coverage in LMICs, and routine diagnostics continue to rely on SSM and clinical diagnosis.

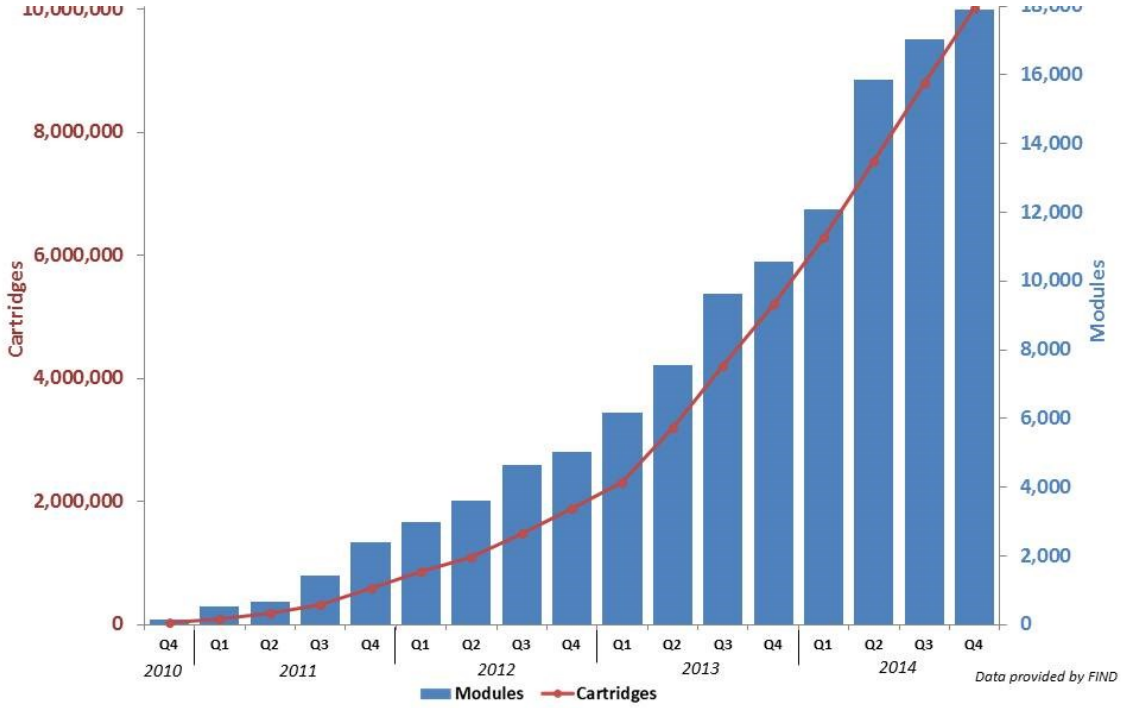


Figure 7. Cumulative number of GeneXpert instrument modules and Xpert MTB/RIF cartridges procured under concessional pricing (48).
(Accessed from: http://who.int/tb/laboratory/GeneXpert_rollout_large.jpg?ua=1)

Another notable post-Xpert phenomenon is the renewed interest in development of novel technological platforms for TB diagnosis which have led to diversified diagnostic portfolio for TB at all levels of the health system (Figure 8) (50). Particularly, many new tests that could potentially be an alternative or can reach much lower laboratory levels than Xpert are now available or emerging from the diagnostic pipeline for assessment of performance, feasibility, and impact in the coming years. The utility and potential impact of these new diagnostics heavily depend on current evidence of impact and limitations of previous generation NAATs (LPA and Xpert) as a routine diagnostic tests for TB. Therefore, it is critical to summarize and evaluate post-implementation impact of previous generation NAATs and identify gaps in research. Furthermore, a thorough assessment of cost-effectiveness of potential implementation strategies of these novel NAATs in settings with and without Xpert will be needed to ensure optimal use of health care resources and placement of these tests.

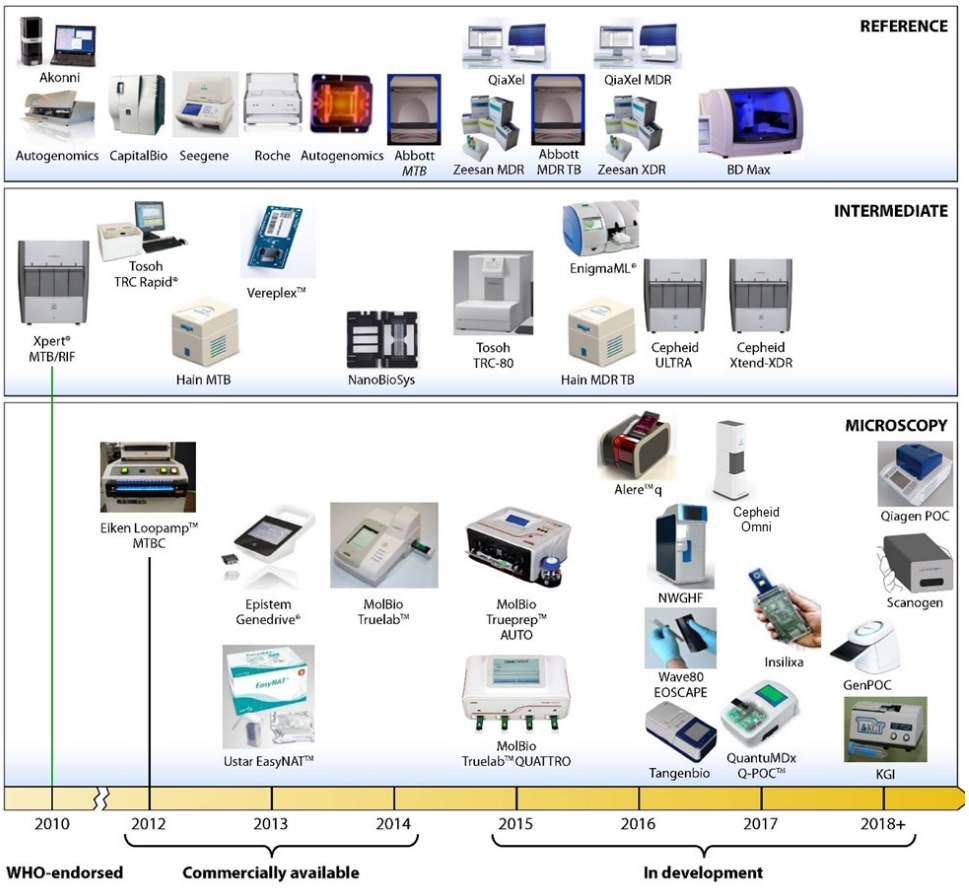


Figure 8. Current and emerging automated, semi-modular or non-integrated TB NAATs; their intended laboratory location and their release or anticipated time to market. (50)

2.3 Evaluation of diagnostic tests

Diagnostic research is a critical component of health technology assessment (HTA), which includes evaluation of test performance, impact and assessment of value for money of diagnostic strategies (51, 52). Likewise, diagnostic research ultimately aims to inform policy and clinical decision-making around the introduction and implementation of new diagnostics.

Unlike evaluation of therapeutic drugs, diagnostic test evaluation does not have well-defined specific phases of diagnostic research (53-56). However, in broader terms, diagnostic research can be described as a process of following study categories (adopted from (57)):

- Assessment of diagnostic technical performance and accuracy (test research)
- Assessment of the impact of diagnostic strategies (additional or replacement of existing diagnostics) using new diagnostic technology on:

① Clinical management (e.g. clinical decision making process)

② Prognosis (e.g. changes in treatment decisions)

③ Health outcome (e.g. patient health quality of life, mortality)

- Synthesis of evidence from multiple studies – systematic review and meta-analysis
- Determination of cost and cost-effectiveness of diagnostic strategies
- Knowledge translation into practice and policy

Therefore, diagnostic research is a comprehensive field of research that includes studying impact elements beyond diagnostic accuracy, such as evidence on clinical decision-making and patient-level outcomes, as well as economic benefits.

2.3.1 TB diagnostics research: the past and present

One of the most valuable assets of a diagnostic test is its ability to differentiate patients with and without disease, generally described by sensitivity and specificity (58). However, test accuracy is, at best, a surrogate for predicting clinical, patient, and public health impact. Likewise, there are a considerable number of factors beyond test performance that can determine potential benefit of a new diagnostic test, such as clinician decisions, operational factors, and ultimately patient characteristics and behaviors. For these reasons, evidence from test accuracy studies are rated as "low-quality" of evidence by the GRADE guidelines, even though these accuracy studies may not be impacted by biases (59).

Prior to the worldwide adaptation of Xpert, TB diagnostic research had prioritized evaluating test accuracy (60). As a result, there was a considerable lack of studies evaluating concepts beyond test accuracy, particularly in a new diagnostic test's contribution to the health care system (e.g. reduction in diagnostic and treatment delays, operational feasibility and improvement), clinical decision making, cost and cost-effectiveness in routine programmatic settings, and on patient outcomes (14).

However, the landscape of diagnostic research in TB has changed considerably in the post-Xpert era, as novel diagnostic tests, including Xpert, possess greater potential to influence factors downstream of diagnosis than previous conventional tests for TB. Subsequently, the policy decision-making processes at global agencies such as WHO have also adopted a two-step approach that incorporates secondary programmatic recommendation made after initial technical recommendation (made mostly based on evidence on test performance) (61). This

secondary recommendation is based on a review of clinical and public health impact after wide adaptation of a new diagnostic test. Likewise, there has been an increased number of studies evaluating impact beyond test accuracy using creative study designs (e.g. hypothetical, pre-post, and before and after) as well as diagnostic randomized control trials (RCT).

2.3.2 Research methods for assessing impact beyond diagnostic accuracy

An ideal design for assessing the impact of any diagnostic test is a diagnostic randomized controlled trial (RCT) (62, 63). However, there are several concerns with a RCT design, in addition to them being expensive, time consuming, and logistically challenging (64). Diagnostic RCTs do not just evaluate a test, they evaluate a strategy or package that includes testing followed by some intervention. Thus, it is not easy to disentangle the efficacy of the test from the efficacy of the intervention (65). Furthermore, it is not easy to capture patient-important outcomes when ethical considerations prevent clinical decision making on the basis of a trial product that is not yet approved for clinical use (as in the case of Xpert) (64). Evidence from RCTs in highly controlled settings may not reflect the real world conditions where diagnostics have to be ultimately deployed (64, 66).

There are alternative methods using quasi-experimental designs, such as before / after studies, to study a test's impact on clinical decision-making and factors that may influence patient-relevant outcomes (e.g. diagnostic and treatment delays). These studies may be implemented retrospectively or prospectively, but given the inflexible nature of data collected retroactively from chart review, prospective studies may be preferred. In some instances, a retrospective study may be the only option when a diagnostic test is implemented solely on the basis of performance characteristics (67). Lack of randomization limits these studies from avoiding internal validity, but can potentially gain external validity through improved generalizability of findings (67). Consequentially, by the nature of the study design, these studies are also subject to a range of potential biases associated with the effects of time. These may include selection bias (e.g. due to policy changes that may directly or indirectly influence patient's eligibility criteria in either period of assessment) and confounding by temporal factors (e.g. change in health care infrastructure, quality of care) (68).

As in some early Xpert studies, a lack of existing policy and guidelines for new diagnostic tests will substantially limit how their impact will be studied during the initial phase of

diagnostic test evaluation. Hence, studies may have to rely on hypothetical designs to assess "potential" impact. In these studies, explicit assumptions must be made of the criteria on which the estimated impact will be assessed. These studies may use a combination of study data, assumptions, and data from other studies to assess changes in clinical management or patient-important outcome 'had the test results been available for doctors to make specific clinical decisions.'

2.3.3 Assessing cost-effectiveness of tuberculosis diagnostic tests

"If resources were infinitely abundant in relation to the demand for them", good decisions in healthcare would only be concerned about interventions that could bring the greatest benefits with minimized harm-doing. However, finite resources mean that explicit choices must be made based on complex considerations that integrate societal values, benefits and harms relative to resource input, and expenditure priorities. Likewise, the primary role of health economic evaluations are to identify resources used against the effects of alternative health interventions to better inform the decision process for efficient and equitable allocation of resources. Therefore, various government and non-government organizations have put a greater emphasis on the use of economic evaluations in the health decision making process (69-71).

Traditionally, decision analytic models have been used to evaluate economic impact of health interventions and have been the primary method for conducting cost-effectiveness analysis for TB diagnostic tests (72). These decision analytic models take account of various probabilities, economic costs, and effectiveness estimates that are associated with diagnostic accuracy: true positive, true negative, false positive, and false negative. Though multiple outcomes can be assessed (e.g. increased case detection, treatment and health outcomes), reference case for cost-effectiveness analysis of health interventions in LMICs generally recommend incremental costs be evaluated against disability-adjusted life years (DALY) averted, which is computed as the incremental cost-effectiveness ratio (ICER) (73). As ICER is a ratio measure, it generally has no meaning unless weighed or compared against some willingness to pay (WTP) threshold such as per-capita gross domestic product (GDP) (74). Likewise, according to the WHO recommendations, a health care intervention is deemed cost-effective if ICER is within the country's per-capita GDP (69).

Decision analysis is a method that involves one or more technical methods such as statistics for deciding among competing alternatives, which the output estimates are based on a hypothetical cohort of target population. Likewise, some key parameters may be largely based on explicit assumptions or borrowed estimates from studies conducted in different settings/countries. Subsequently, cost-effectiveness analysis models used in TB research are subject to several important limitations. First, the models primarily rely on diagnostic test performance estimates to project outcomes important to clinical management, patients, and ultimately to public health. Additionally, they are subject to same type of limitations discussed in earlier sections (e.g., limitations of using diagnostic accuracy as surrogate outcomes beyond diagnosis). Existing models are based on diagnostic algorithms and do not factor operational characteristics such as diagnostic coverage scenarios. Furthermore, a systematic review of cost-effectiveness analysis of Interferon Gamma Release Assays (IGRA) for the diagnosis of latent TB infection (LTBI) has found that there are considerable methodologic differences amongst the studies included, particularly on how cost parameters were evaluated in the model. Many studies lacked proper economic assessment of costs associated interventions evaluated in the study, neglecting important opportunity costs that may be associated with implementation and operation of an intervention (75). This is particularly concerning as cost is one of the two most important components of cost-effectiveness analysis, and the result of the inclusion or exclusion of certain cost parameters may not truly reflect the realities of a given intervention (i.e. costs of implementation and scale-up of new diagnostics can divert resources from other potentially useful interventions) (72). Therefore, it is critical implement good costing practices when evaluating cost-effectiveness of a public health intervention.

Chapter 3: Impact of molecular tuberculosis diagnostics on improving time delays in diagnosis and treatment of tuberculosis: a systematic review

3.1 Preface

Line probe assays (LPA) and Xpert MTB/RIF® (Xpert) are commercial nucleic acid amplification tests (NAATs) that have been endorsed by WHO on the basis of anticipated improvement in accurate and timely diagnosis of tuberculosis (TB). Considerable efforts and commitment have been made by governments of high TB burden countries to scale-up these NAATs to improve diagnostic capacity and infrastructure of TB diagnosis. Subsequently, increasing research efforts have been made to study impact beyond diagnosis, particularly in evaluating the impact on diagnostic and treatment delays.

Existing systematic reviews of these tests have been limited to summarizing diagnostic accuracy. Furthermore, while there are systematic reviews assessing patient, diagnostic and treatment delays, no systematic review exists that evaluates impact of diagnostic test and interventions on reducing time delays associated with diagnosis and treatment of TB. Also, there is considerable inconsistency in how time delay components are defined and assessed. These inconsistencies range from inclusion and exclusion of time delay components to inappropriate statistical assessment of time data. These discrepancies can complicate and obstruct direct comparison across multiple studies, limiting generalizability of the study findings.

We used this study to critically summarize and quantitatively assess actual impact of WHO-approved NAATs on reducing time delays in diagnosis and treatment of TB. Moreover, we propose an improved framework to classify critical time delay components. We also used this opportunity to develop a criteria to assess quality of time delay estimates to highlight methodological areas of concern in studying time delays, which is an area that was not adequately addressed in previous systematic reviews studying delays in diagnosis and treatment of TB.

3.2 Manuscript

Impact of molecular diagnostic tests on TB diagnostic and treatment delays: a systematic review and meta-analysis

Hojoon Sohn[1,2], Samuel G. Schumacher[1,2], Andrea Benedetti[1,2], Zhi Zhen Qin[1,2], Genevieve Gore[3] Madhukar Pai[1,2]

Affiliations

[1] McGill University Department of Epidemiology & Biostatistics

[2] McGill International TB Centre

[3]McGill University, Schulich Library of Science and Engineering

Abstract

Rationale: Delays in diagnosis and treatment of tuberculosis (TB) are critical obstacles for TB control. The World Health Organization (WHO) endorsement of Nucleic Acid Amplification Tests (NAATs) such as Xpert MTB/RIF and line probe assays (LPA) have resulted in unprecedented efforts to scale-up diagnostic capacities in high TB burden countries (HBCs), and increasing interest in research focusing on patient impact of NAATs.

Objective: Critically summarize and assess the literature concerning comparative impact of NAATs on diagnostic and therapeutic delays compared to the standard diagnostic methods for drug susceptible and drug-resistant TB using both narrative and quantitative methods.

Methods: Prior to the review, a conceptual framework for defining time delays from the on-set of disease symptoms to initiation of anti-TB treatment was developed. We searched MEDLINE, EMBASE, Web of Knowledge and Cochrane CENTRAL databases. For numeric assessment of the impact on delays, we calculated absolute mean reduction of diagnostic and therapeutic delays for each study and pooled them using random effects meta-analysis.

Main Results: A total of 39 eligible studies were included in this review. We found considerable methodological inconsistencies in how time delays were estimated across the studies. Use of Xpert reduced the delays by 2.38 days (95% CI -0.09, 4.85) for diagnosis and 16.54 days (95% CI 6.73, 26.35) for treatment of drug-susceptible TB relative to diagnosis and treatment based on sputum smear microscopy. LPA reduced delays by 46.57 days (95% CI 28.89, 64.23) and 62.48 days (95% CI 27.72, 97.24) for diagnosis and treatment of drug-resistant TB respectively relative to conventional culture drug-susceptibility testing. The magnitude and significances of these effect estimates varied considerably in sub-group analyses, as studies includes were highly heterogeneous.

Conclusions: Our meta-analysis showed that use of NAATs reduced both diagnostic and therapeutic delays. Future studies assessing impact of novel diagnostic tests on time delays

should use standardized measures of time delays, and account for factors that might influence time delays.

Introduction

Despite a steady decline, tuberculosis (TB) remains a major global public health concern with more than 9.6 million TB cases and 1.5 million deaths reported in 2014 (2). Only a third of the 480,000 estimated new cases of Multi-drug Resistant TB (MDR-TB) were diagnosed and approximately 30% of diagnosed patients failed to receive proper treatment (2). While reducing time delays to diagnosis and treatment are important obstacles to the control of the global TB epidemic (76, 77), high TB burden countries (HBCs) continue to rely on century-old diagnostics techniques that compromise accurate and timely diagnosis of TB.

Line probe assays (LPA) and Xpert MTB/RIF® (Xpert) are commercial nucleic acid amplification tests (NAATs) that have been shown to have good diagnostic accuracy with the capacity to diagnose drug susceptible (DS) and resistant (DR) TB within 1-2 days of sample processing (42, 78). Anticipating improvements in accurate and timely TB diagnosis, these NAATs were endorsed by the World Health Organization (WHO) (11, 12, 45). Since the WHO policy, unprecedented efforts have been made to scale-up Xpert and LPA (13-15). With solid evidence on their test accuracy, research has increasingly focused on studying their actual clinical impact (14, 61, 79-82). While there are systematic reviews on the diagnostic accuracy of Xpert and LPAs (42, 83, 84), and others that separately describe diagnostic and treatment delays experienced by TB patients (6, 32, 85), no systematic review has summarized the impact of NAATs on reduction of time delays in diagnosis and treatment of TB.

Therefore, the main objective of our systematic review was to critically summarize, using both narrative and quantitative methods, the available evidence of the impact of NAATs on diagnostic and therapeutic delays compared to that of the standard of care for DS and DR-TB. We also highlighted methodological areas of concern in assessing time delays, an aspect that

has not been adequately addressed in previous systematic reviews of diagnostic delays in TB. Our review will contribute to on-going efforts in improving the diagnostic capacity and effective operationalization of diagnosis and treatment of TB in HBCs.

Methods

Study selection criteria and operational definitions

Prior to the review, a study protocol was developed based on the standard procedures of the Cochrane Collaboration (86) and the Preferred Reporting Items for Systematic Reviews and Meta Analyses (PRISMA) statement and checklist (87). We then developed a conceptual framework for classification of time delay components that are critical for operational and clinical impact of diagnosing and treating TB disease. Essential time delay components and definitions of each time delay categories were developed based on an existing framework (32, 33) and are illustrated in Figure 1. This framework provides a standardized and structural guidance in assessing time delays reported in the studies included in this review.

Our review focused on the impact of the WHO-approved TB NAATs, specifically the GenoType MTBDR*plus* (Hain Life Sciences GmBH, Nehren, Germany) and Inno-LiPA RifTB (Fujirebio Europe N.V, Gent, Belgium) – both referred as LPA here on, and Xpert® MTB/RIF assay (Cepheid, Sunnyvale, CA, USA) – referred as Xpert here on. LPAs are currently approved for testing from direct sputum (smear positive) or culture isolate (smear negative patients) who are suspected of MDR-TB (11). Xpert is currently recommended as an initial test in those who are at risk of MDR-TB or HIV associated TB and as an add-on test for those patients who are smear negative, but not at risk of MDR-TB or HIV associated TB (conditional recommendation with acknowledging resource constraints) (45).

While a variety of products with similar technological platform are available in the market, these tests are the only NAATs approved by the WHO, which has led to rapid scale-up and implementation at the global level. We excluded studies focused only on childhood and extra-pulmonary TB, since these conditions are difficult to diagnose and time delays are very challenging to assess. We included only peer-reviewed studies that assessed time delays in the process of diagnosis and treatment of TB and MDR-TB, and that used NAATs as an index test. We did not restrict our studies based on region, setting, years, language, or types of study design, but only included studies with primary data, thus excluding reviews and modeling studies. Studies only reporting 'run-time' or turn-around-time of the test (e.g. "2 hours to run" Xpert test) were excluded from our review.

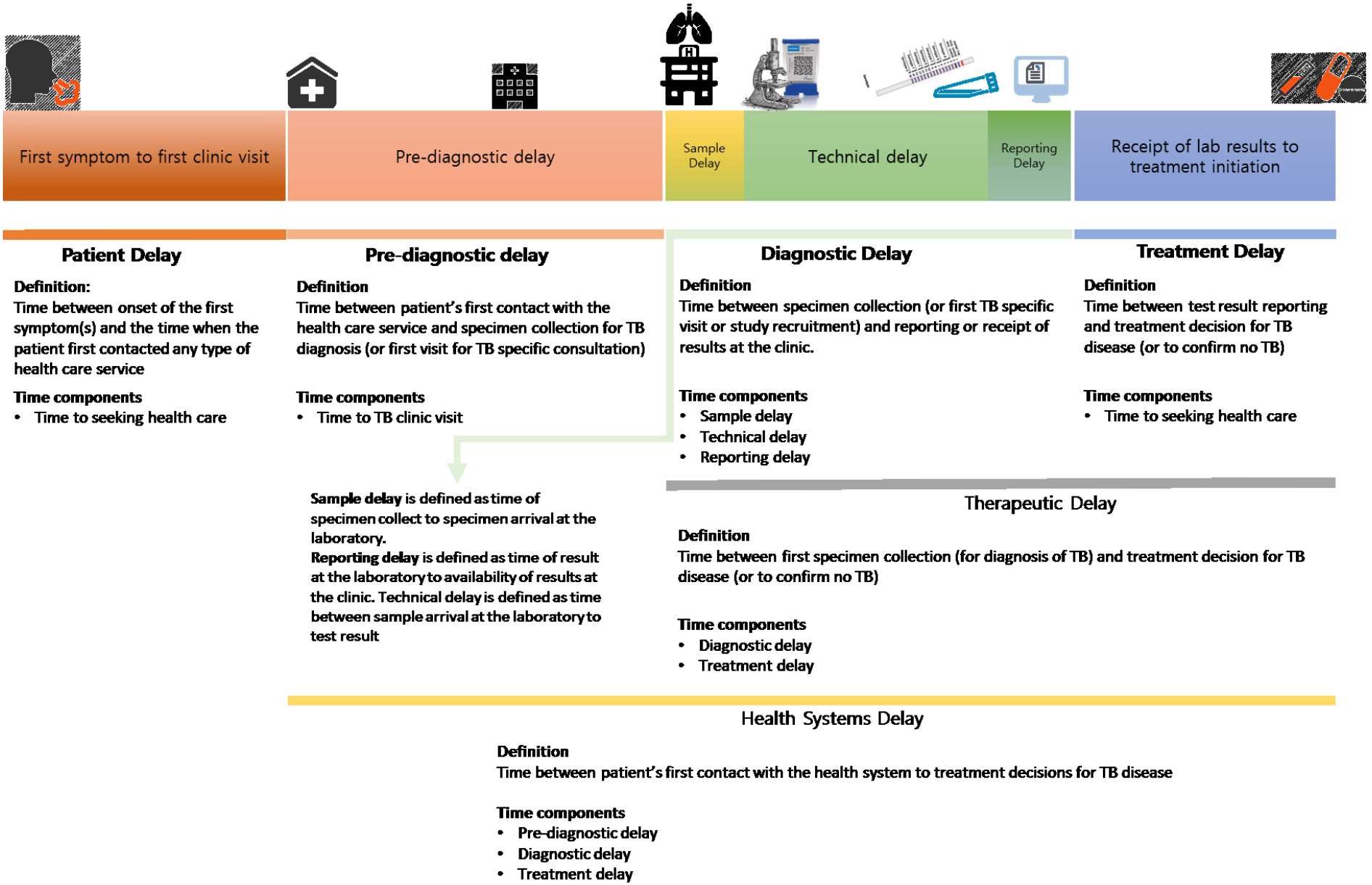


Figure 1. Conceptual framework of time delay components in diagnosis and treatment of tuberculosis

Search strategy, study selection, and data extraction

The complete electronic search strategy is available in Appendix A. Eligible studies were identified from MEDLINE, EMBASE, Web of Knowledge and Cochrane CENTRAL databases through searches incorporating terms associated with time (e.g. “delay”, “time”, “turnaround time”, “time to treatment”) in addition to specific terms to limit search within LPA and Xpert for diagnosis of adult pulmonary TB and MDR-TB (i.e. excluded childhood and extra-pulmonary TB). The first search was conducted on January 10, 2014. With an extensive list of Xpert related studies published in 2014, we conducted an update search on January 31st, 2015. Ongoing trials, if applicable, were identified based on search results from the metaRegister of Controlled Trials (mRCT) and the WHO International Clinical Trials Registry Platform. References of included articles and previous systematic reviews focusing on diagnostic accuracy of NAATs were consulted, and experts in the fields of TB diagnostics were contacted to identify additional studies not included in the database search.

The list of titles and abstracts were screened independently by two reviewers (SGS, ZQ – screen 1, HS, SGS – screen 2). Any potentially eligible studies identified from the two screenings were selected for full-text review and independently assessed for inclusion, based on the study protocol, by two independent reviewers (HS, SGS). A full list of excluded studies with reason for exclusion was updated once discrepancies in inclusion and exclusion were resolved by the two reviewers.

A standardized data extraction form was developed using Google Forms (Google Inc., Mountain View, CA, USA) to help reduce errors during data entry (88). Six of the 39 studies included in this review were used to pilot test the data extraction form and the data parameters were independently extracted by the two reviewers (HS, SGS). The form was subsequently revised to improve the quality of data extracted concerning study design, key

epidemiologic and contextual factors, time delay, and quality assessment. One reviewer (HS) then extracted data for the remaining studies and a second reviewer (SGS) examined all of the extracted items for any discrepant data. We abstracted numeric (median or mean) data on time delays for both the index tests (Xpert or LPA) and the comparator tests (conventional TB diagnostics) according to our operational definitions, up to the point of treatment initiation; some studies reported time or time relevant data beyond treatment initiation such as time to culture conversion. Units of time were converted into number of 'days' if reported in other units (e.g. months, weeks, or hours) (89-91).

Quality assessment of time delay estimates

Unlike quality assessment tools for diagnostic accuracy studies (92), there currently is no established method or checklist(s) that can be used to assess the quality of studies investigating time delays or time to event study outcomes. Therefore, we developed a matrix of key methodologic and contextual information necessary to determine the usefulness and comparability of the time delay reported. These included 1) provision of clear definition of measuring time delay ("delay definition"); 2) empirical or hypothetical assessment of time delay ("empirical vs. hypothetical"); 3) assessment of complete diagnostic or therapeutic time delay ("complete impact"); 4) use of proper statistical methods for measurement and reporting of time delay ("measurement and reporting"); and 5) proper assessment and adjustment of factors that could influence time delay outcomes ("risk factors"). Each category was assigned values of 0 or 1 where 1 was considered positive assessment of the quality.

For "delay definition", we assigned value of 1 if authors provided clear start and end time definitions in assessing time delay anywhere in the text. For instance, if a study reported time

to treatment without indicating starting time or left it ambiguous, this category would take a value of 0. Studies assessing time delays based on hypothetical assessment of 'would have been' time of clinical action took the value of 0 for "empirical vs. hypothetical". If studies reviewed health records or individual patient time stamp of events, a value of 1 would be assigned. If studies only assessed technical delay or just one component of diagnostic or therapeutic delay, "complete impact" was assigned a value of 0, 1 otherwise.

For "measurement and reporting", we considered median time delay with interquartile range (IQR) most adequate with reporting time delay, as time duration does not generally follow normal distribution. Thus, studies reporting medians with IQRs, with p-value statistics were assigned a value of 1. For those studies reporting mean/average time with 95% confidence interval supported by proper methodological indication (i.e. normal distribution of time in the study), these studies also took a value of 1 for this category. Otherwise, those studies lacking these credentials (including studies lacking p-value statistics) were assigned a value of 0. If the study reported or properly adjusted for factors that could potentially influence time delays in diagnosis and treatment of TB using NAATs, they were assigned the value of 1 for "risk factors" (randomized control studies were assigned 1 as they will generally produce balance in underlying risk factors for delays), 0 otherwise. Scores from each category were summed to represent 'quality' of time delay reported, with 0 being the weakest and 5 being the best.

Data synthesis

We first conducted a comprehensive comparative and qualitative assessment of the time delays reported using descriptive and narrative summaries of the studies included. We calculated overall medians and IQRs of diagnostic and therapeutic delay for each diagnostic test (NAATs vs. smear, culture, and culture DST) from the respective medians reported by

individual studies, which were graphically summarized in box-plots. The Kruskal-Wallis test was used to test for statistical significance of the differences in the medians (93).

Next, meta-analyses were conducted to assess the absolute reduction in time delay in diagnosis and treatment of TB using NAATs. Time to event and time delay data are non-normally distributed variables that are reported with medians and IQRs. This posed problems in conducting meta-analysis as 1) conventional meta-analytic methods for means require assumptions of normal distribution of the data and 2) there is lack of well-established non-parametric meta-analysis methods to summarize time delays. To cope with these challenges, we first converted time delay estimates reported in medians (with IQRs) into mean delay with respective 95% CI using the methods reported by Wan et al. (94). As units of delay measurements (days) were uniform across all studies, we then estimated the effect size as the raw mean differences in time between NAATs and conventional tests for both diagnostic and therapeutic delays (95). These effect size estimates of raw differences were meta-analyzed using a random effects model, since we expected that the observed variations in the delays between studies are likely to be caused by more than random chance (96, 97).

Heterogeneity was evaluated based on the I-squared statistic where a value greater than 75% is considered to be highly heterogeneous (98, 99). Sub-group analyses were conducted to identify possible sources of heterogeneity and to assess key factors that can variably influence magnitude of our effect size estimate. Given that the majority of the studies included are observational and that the effect estimates are time delays, publication bias due to unpublished non-significant studies was considered unlikely and therefore was not formally assessed in this review. Data were managed and analyzed using Microsoft Excel 7.0 (Microsoft Corporation, Redmond, WA, USA) and R (R Foundation for Statistical Computing, Vienna, Austria).

Results

Search results

The complete process from screening to identifying studies included in this review is illustrated in Figure 2. After removal of duplicates, we screened a total of 7,995 titles and abstracts, which identified 107 studies eligible for full text review. A total of 71 articles were excluded (major reasons for exclusion listed in Figure 2). A total of 39studies were included in this review, including 3 identified after second screening. Two studies (91, 100) did not report time delays for the comparator, but were included in our review for the comparative assessment of time delays within the index test.

Description of included studies

Of the 39 studies included (Table 1 and 2) in this review, 24 (62%) were conducted in World Bank-classified Low and Middle Income Countries (LMICs), a measure based on gross national income (GNI) in 2015. Of these studies, 15 (38%) were conducted in African region, where the majority of the studies in this region were from South Africa (n = 13) either as part of a multi-center or independent study. Xpert was studied as the index test or intervention in 24 studies (62%), with 2 investigating Xpert exclusively as a test to diagnose and treat DR-TB (both conducted in South Africa).

A total of 28 studies reported HIV prevalence, of which half (n =14) were in settings with HIV prevalence greater than 50% (13 studies for Xpert, 1 for LPA) and 13% (5 studies for Xpert) of studies were done exclusively in those infected with HIV. There were 12 studies reporting the rate of empiric treatment for TB (all from Xpert studies) of which 5 reported – 4 from South Africa and 1 from Zimbabwe – rates greater than 40% (51-69%).

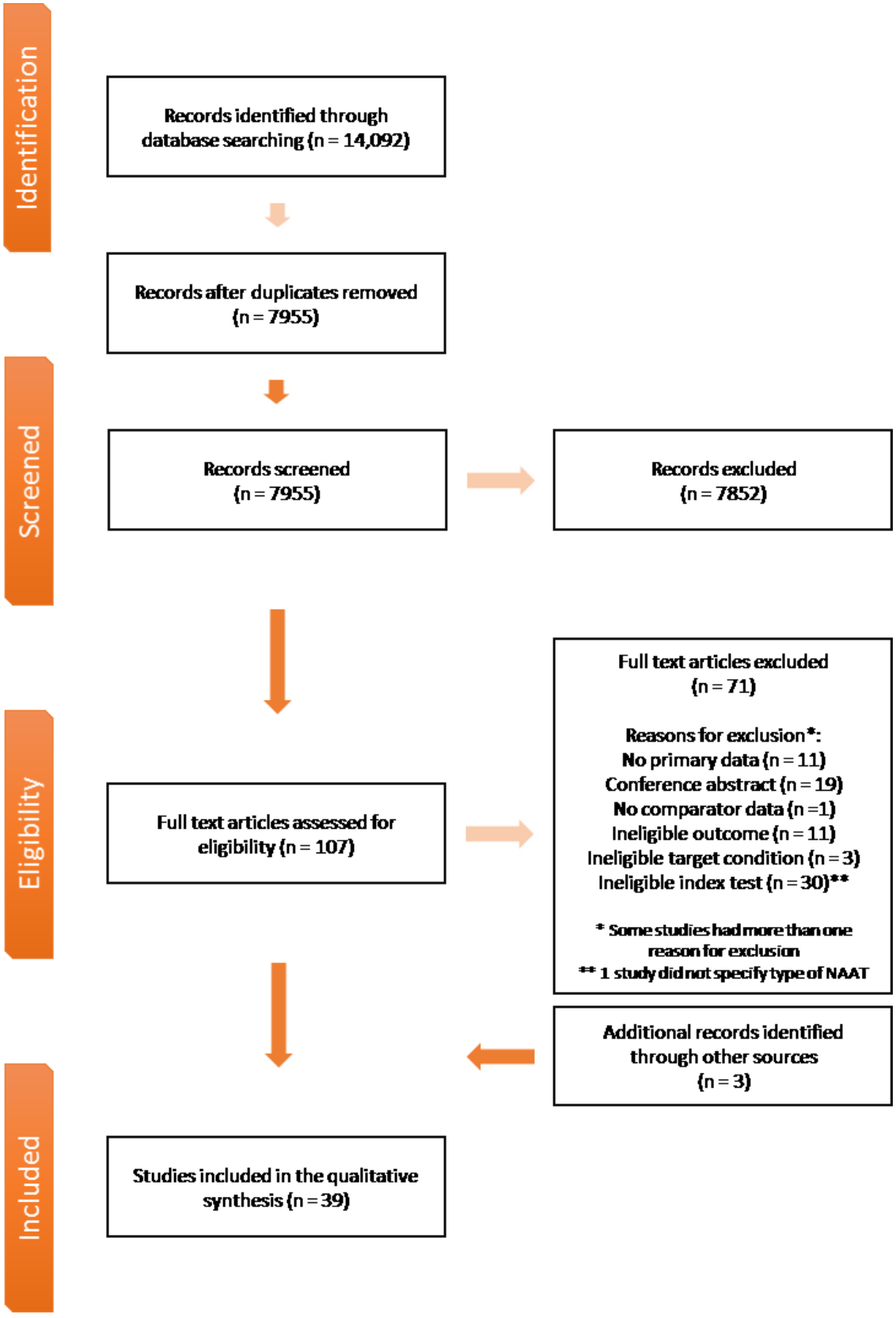


Figure 2. Study selection

	Study Design	Country	Setting	Author's time delay terminology	Diagnostic Delay (Median, IQR)			Therapeutic Delay (Median, IQR)			Commencement time
					Index	Comparator	p-value	Index	Comparator	p-value	
Theron (2014) (66)	Ind. RCT	Multiple*	Urban Clinic POCT	Time to diagnosis and Time to treatment	81% diagnosed on same day	43% diagnosed on same day	< 0.0001	Same day (0-3)	1 day (0-4)	< 0.0001	Enrolment+
Mupfumi (2014) (101)	Ind. RCT	Zimbabwe	Urban Hospital Off-site	Time to diagnosis and Time to treatment initiation,	2 days (0-13)	6 days (1-25)	0.07	5 days (3-13)	8 days (3-23)	0.25	Clinical presentation
Cox (2014) (102)	Parallel Cl. RCT	South Africa	Urban Clinic On-site	Time to TB treatment	NR	NR	NR	4 days (2-7)	8 days (2-27)	NR	Enrolment+
Durovni (2014) (103)	St.-we Cl. RCT	Brazil	Urban Clinic Off-site	Time to positive result, Time to treatment initiation, measured from specimen collection	7.3 days (3.4-9.0)	7.5 days (4.9-10.0)	0.51	8.1 days (5.4-9.3)	11.4 days (8.5-14.5)	0.04	Specimen collection
Churchyard (2015) (104)	Parallel cRCT	South Africa	Urban/Rural Clinic Off-site	Time to starting treatment	NR	NR	NR	7 days	10 days	NR	Enrolment+
Boehme (2011) (105)	Pre/Post	Multiple**	Urban Mixed On-site	**Time to detection**, Time to Reporting, and Time to treatment	1 day (0-2)	**Smear:** 2 days (2-3) **Culture:** 58 days (42-62)	NR	5 days≠ (2-8)	56 days (39-81)	NR	Collection of first sputum
Yoon (2012) (106)	Pre/Post	Uganda	Urban Hospital On-site	**Time to detection**, Time-to-TB treatment	Same day (0-2)	1 day (0-26)	<0.001	6 days≠ (1-61)	7 days (3-53)	0.06	Enrolment
Chaisson (2014) (90)	Hypothetical	USA	Urban Hospita On-site	Time to specimen collection, Time to result, Total time in isolation	1 day (0-2)	2 days (1-4)	NR	< 2 days (1-2) Hypothetical	35 days (5-65)	NR	Time from admission
Sohn (2014) (107)	Hypothetical	Canada	Urban Hospital POCT	Time to reporting, time to treatment initiation	1 day (0-4)	**Smear:** 1 day (1-2) **Culture:** 21.5 days	NR	Hypothetically reduce by 12 days (4-23) in smear negative TB patient	26 days (4-30)	NR	Time from first sample
Lippincott (2014) (89)	Hypothetical	USA	Urban Hospital On-site	AII initiation to sample at laboratory, Laboratory processing time, AII duration	< 1 day (0-1)	> 1 day (1-2)	< 0.04	< 1 day (0.7-1.33) Hypotehtical	3 days (2-4)	< 0.04	Time from AII initiation
Davis (2014) (108)	Hypothetical	USA	Urban Hospital On-site	Duration of over-treatment	NR	NR	NR	1 day (1-3) Hypothetical	46 days (45-49)	NR	Enrolment+
Balcells (2012) (109)	Observational	Chile	Urban Hospital Off-site	Time reduction to initiation of proper anti-TB treatment	NR	NR	NR	"Median time of 14 days earlier" vs. waiting for culture result	-	NR	Enrolment+
Van Rie (2013) (110)	Observational	South Africa	Urban Clinic POCT	Time to treatment initiation	NR	NR	NR	same day	13 days (10-20)	< 0.001	Enrolment (time of 3rd sputum)

Hanrahan (2013) (111)	Observational	South Africa	Urban Clinic POCT	Time to treatment initiation	NR	NR	NR	**1st Xpert positive:** same day	**1st Xpert negative** - Empiric TB: 14 days (5-35) - Culture +: 114 days (28-180)	NR	Baseline visit (initial Xpert test performed)
Kwak (2013) (112)	Observational	South Korea	Urban Hospital On-site	**Turn around time**, Time to confirmation of receipt of results, Time to anti-TB treatment	6 days (3-7)	**Smear:** 12 days (7-19.25) **Culture:** 38.5 (35.75-50.25)	< 0.001	7 days (4-9)	21 days (7.33.5)	< 0.001	Request of diagnostic test
Buchelli-Ramirez (2014) (113)	Observational	Spain	Urban Hospital On-site	System-related treatment delay	NR	NR	NR	**Smear negative:** 15.5 (1.25, 28.7) ¥	42 days (22, 61) ¥	NR	Patient's first contact with the health care system (first consultation)
Cohen (2014) (114)	Observational	South Africa	Urban Hospital Off-site	Total diagnostic time (3 sub categories)	6.3 days (5.3-8.1)	3.3 days (2.1-5.2)	< 0.001	NR	NR	NR	Sputum collection
Kim, SY (2012) (115)	Observational	South Korea	Urban Hospital On-site	Turn around time (laboratory)	< 1 day	**Smear -, Culture +:** 19 days (range 9-48)	NR	NR	NR	NR	Submission of sample(s) +
Kim, CH (2014) (116)	Observational	South Korea	Urban Hospital On-site	Turn around time	< 1 day (0-4)	**MTB nested PCR:** 4 days (1-11)	< 0.001	NR	NR	NR	Submission of sample(s) +
Auld (2014) (100)	Observational	Cambodia	Urban Hospital Off-site	Turn around time	1 day (0-7)	NR	NR	NR	NR	NR	Submission of samples
Omrani (2014) (117)	Observational	Saudi Arabia	Urban Hospital On-site	Time to reporting, Time to anti-TB therapy	1 day ("IQR 4 days")	**Smear:** 1 day ("IQR 1 day") **Culture:** 44 days ("IQR 30 days")	NR	Same day ("IQR 3 days")	**Smear:** Same day ("IQR 1 day") **Culture:** 22 days ("IQR 21 days")	NR	Sample collection

Table 1. Study characteristics and time delays reported for diagnosis and treatment of drug susceptible TB

* South Africa, Zimbabwe, Zambia, and Tanzania. ** South Africa, Peru, India, Azerbaijan, Philippines, and Uganda
+ estimated based on study design
≠ in smear negatives
¥ reported as mean & 95% confidence interval (for Jacobson, mean and 95% confidence interval is calculated from median and IQR reported (94) for specimen delay and technical delay)
Author defined four categories of time between identification of patients suspected of MDR-TB to initiation of MDR-TB treatment. Provided total delay and time delay for each time categories (identification to specimen arrival at laboratory, laboratory processing time, laboratory result to patient receiving results, receipt of results to initiation of treatment)
$ 19 out of 38 Xpert negative with chest X-ray started on treatment

Ind. RCT=Individually Randomized Controlled Trial. Cl. RCT=Cluster Randomized Controlled Trial. St.-we. Cl. RCT=Stepped-wedge Cluster Randomized Controlled Trial. Pre/post=Pre/post implementation study. Hypothetical=Single-cohort hypothetical study. Observational= Single-cohort observational study. AII=Airborne Infection Isolation. POCT=Point-of-care testing. TB=Tuberculosis. MDR=Multidrug-resistant tuberculosis. NR=not reported.

	Study Design	Country	Setting	Author's time delay terminology	Diagnostic Delay (Median, IQR)			Therapeutic Delay (Median, IQR)			Commencement time
					Index (LPA)	Comparator (Culture)	p-value	Index	Comparator	p-value	
Naidoo (2014) (118)	Pre/Post	South Africa	Urban Clinic Off-site	**Time to detection**, Time to commencement of treatment (action delay)	**Xpert:** Median < 1 day 95% CI (<1, 17)¥	**LPA:** Median 24 days 95% CI (18, 33) ¥	< 0.001	**Xpert:** Median 17 days 95% CI (7, 36)	**LPA:** Median 43 days 95% CI (30, 64)	< 0.01	Date of sputum collection
Cox (2015) (119)	Pre/Post	South Africa	Urban Clinic Off-site	Time to treatment initiation	NR	NR	NR	**Xpert:** 2012: 13 days (6-35) 2013: 8 days (5-25)	**LPA (2011):** 28 days (16-40) Culture (2006): 71 days (49-134)	NR	Date of sputum collection
Dlamini-Mvelase (2014) (91)	Observational	South Africa	Urban Hospital On-site	Time to treatment (% of patient on Tx. by time category)	NR	NR	NR	Estimated 50% of patients of receive treatment within 20 days of Xpert	NR	NR	Registry of Xpert test
Skenders (2011) (120)	Pre/Post	Latvia	Urban Hospital Unclear	Turn around time, treatment start, culture conversion, and final treatment outcome	*Direct:* 10 days (6.6-13.4) *Culture Isolates:* 34.2 days (25.1-43.3)	22.4 days (16.8-28.1)	< 0.001	14 days (7-22)	40 days (23-67)	< 0.0001	Hospital admission
Hanrahan (2012) (121)	Pre/Post	South Africa	Unclear Clinic Off-site	Laboratory turn around time**,** Time to MDR-TB treatment	*Direct:* 26 days (11-52) *Culture Isolates:* 29 days (22-43)	52 days (41-77)	0.008	*Direct:* 54 days (31-66) *Culture Isolates:* 73.5 days (43-94) Overall: 62 days (32-86)	78 days (52-93)	0.045 (for overall)	Date of specimen collection
Jacobson (2012) (122)	Pre/Post	South Africa	Rural Hospital Off-site	Time delay: specimen collection to specimen arrived at laboratory, laboratory processing time, clinic notification of DST result to start MDR treatment	*Culture isolates:* 29 days (19, 40) ¥	58 days (43, 73) ¥	NR	*Culture isolates:* 55 days (37.5-78)	80 days (62-100)	NR	Date of sputum collection
Kipiani (2014) (123)	Pre/Post	Georgia	Urban Hospital Off-site	Time to treatment initiation	NR	NR	NR	*Direct:* 18.2 days (11-24)	83.9 days (56-106)	< 0.01	Date of first sputum collection
Singla (2014) (124)	Pre/Post	India	Urban Hospital On-site	See below #	*Direct:* 28 days (7-49) ¥	138 days (91, 181) ¥	< 0.0001	*Direct:* 38 days (30-79)	157 days (127-200)	< 0.001	Identification of patients suspected of MDR-TB
Lyu (2013) (125)	Observational	South Korea	Urban Hospital Off-site	Time to reporting, Time to anti-TB treatment initiation	*Direct* 12.7 days (8-17) *Culture isolate* 32.6 days (14-48)	83 days (68-92)	< 0.01	**Treatment delay€** *Direct:* 19.8 days (9-26) *Culture isolate:* 43.2 days (31-52)	88.9 days (64-111)	< 0.05	Specimen request
Raizada (2014) (126)	Observational	India	Urban Mixed Off-site	Turn around testing time	*Direct:* 11 days (range 1-76)	87 days (42-208)	NR	NR	NR	NR	Specimen collection
Barnard (2008) (127)	Observational	South Africa	Urban Mixed	Turn around time	*Direct:* 2 days ¥	42 days (32,51) ¥	NR	NR	NR	NR	Specimen at laboratory

			Off-site								
Anek-vorapong (2012) (128)	Observational	Thailand	Urban Mixed Off-site	Turn around time	*Direct:* 3 days (2-5) *Culture isolates:* 16 days (14-21)	25 days (21-29)	NR	NR	NR	NR	Specimen at laboratory
Tukvadze (2012) (129)	Observational	Georgia	Urban Mixed Off-site	Turn around time	*Direct:* 4.2 days (2.4-6) ¥	*Liquid culture:* 21.6 days (12.3, 30.9) ¥ *Solid culture:* 67.5 (52.5, 82.5) ¥	NR	NR	NR	NR	Sputum collection
Seoudi (2012) (130)	Observational	United Kingdom	Urban Hospital Off-site	Time to diagnosis	**** Days saved before culture result** *Direct:* 25. 3 days (5.94, 44.66) *Culture isolates:* 47.57 days (21.25, 73.89)		NR	NR	NR	NR	Receiving sample at laboratory
Chryssanthou (2011) (131)	Observational	Sweden	Urban Mixed Off-site	Laboratory processing time	*Direct:* 7 days (range 1-16)	21 days (13-78)	NR	NR	NR	NR	Specimen arrival at laboratory
Gauthier (2014) (132)	Observational	Haiti	Urban Mixed Off-site	Turn around time	*Direct:* 7.5 days (range 6.5-8.5)	54 days (range 13-78)	NR	NR	NR	NR	Specimen arrival at laboratory
Martinez-L (2014) (133)	Observational	Spain	Urban Mixed Off-site	Time to detection	8 days (NR)	41.5 days	NR	NR	NR	NR	Specimen arrival at laboratory
Singhal (2014) (134)	Observational	India	Urban Mixed Off-site	Turn around time	*Direct:* 5.2 days (NR) *Culture isolate:* 23.4 days	84 (63-84)	NR	NR	NR	NR	Specimen arrival at laboratory

Table 2. Study characteristics and time delays reported for diagnosis and treatment of drug resistant TB

* South Africa, Zimbabwe, Zambia, and Tanzania. ** South Africa, Peru, India, Azerbaijan, Philippines, and Uganda
+ estimated based on study design
≠ in smear negatives

¥ reported as mean & 95% confidence interval (for Jacobson, mean and 95% confidence interval is calculated from median and IQR reported (94) for specimen delay and technical delay)

Author defined four categories of time between identification of patients suspected of MDR-TB to initiation of MDR-TB treatment. Provided total delay and time delay for each time categories (identification to specimen arrival at laboratory, laboratory processing time, laboratory result to patient receiving results, receipt of results to initiation of treatment)
$ 19 out of 38 Xpert negative with chest X-ray started on treatment

€ Time between results reporting and treatment initiation

Ind. RCT=Individually Randomized Controlled Trial. Cl. RCT=Cluster Randomized Controlled Trial. St.-we. Cl. RCT=Stepped-wedge Cluster Randomized Controlled Trial. Pre/post=Pre/post implementation study. Hypothetical=Single-cohort hypothetical study. Observational= Single-cohort observational study. AII=Airborne Infection Isolation. TB=Tuberculosis. MDR=Multidrug-resistant tuberculosis. NR=not reported.

Of the studies evaluating Xpert, 19 studies (79%) reported using chest radiography (CXR) for diagnosing TB. 4 studies (16%) implemented Xpert as a point of care testing (POCT) program and 13 (52%) implemented Xpert on-site, within walking distance of a primary care clinic or within the hospital laboratory. POCT program were generally defined by each study as Xpert testing performed by non-laboratory personnel within the TB clinic. LPA was implemented within the study site (hospital laboratory) in 7 out of 15 studies (46%) evaluating LPA.

Quality of included studies

Figure 3 graphically summarizes the distribution of the included studies' overall quality scores and the frequency of studies receiving a score of '1' for each quality assessment category. The quality of reported time delay estimates widely varied across the 39 studies included in this review. Using our quality assessment scale, only two studies (5%) received a complete score of 5. Twenty nine studies (76%) received a score of 3 or lower.

Major areas of concern regarding quality were 1) lack of proper assessment and adjustment of risk factors affecting time delays (77%), 2) improper or lack of statistical assessment of time delay (64%), and 3) lack of clear author-defined time delays (46%). Many of the laboratory-based (28%) studies reported partial diagnostic delay, only reporting technical delays. Four studies (10 %) used hypothetical time points to determine the time impact of the index test on therapeutic delay. A graphical summary of overall quality assessment score for each individual studies is available in Appendix B.

Definitions of time delays and study designs

When classifying reported time delays according to our operational definition and by study design (Figure 4), only one study reported all sub-components of time delay and provided overall diagnostic and therapeutic delay (124). Only one other study (114) evaluated detailed

sub-components of complete diagnostic delay. A total of 16 (41%) studies (66, 89, 90, 101, 103, 106, 107, 112, 117, 118, 120-122, 124, 135) reported both diagnostic and therapeutic delays whereas 8 (21%) studies (127-134) only reported technical delays (i.e. laboratory turn around or processing time). A majority of studies (74%, n = 29) measured time delay either from enrolment or first specimen collection. The remaining 10 studies (26%) measured the delay from the time specimens arrived at the laboratory. All of the studies evaluating therapeutic delay used anti-TB treatment initiation time as the end point in evaluating therapeutic delay. Two studies (120, 123) also evaluated time delay beyond treatment initiation (time to culture conversion, mean days in DS-TB ward, days to initiate contact tracing, mean total hospitalization). None of the studies included in this review investigated patient or pre-diagnostic delays.

A total of 5 (13%) studies employed a randomized control (RCT) study design, two (2) of which were cluster RCTs. 9 (23%) studies were quasi-experimental studies using pre/post implementation study designs. 25 (64%) studies were observational studies where 4 (10%) of these studies assessed the potential impact of Xpert on time delay using a hypothetical design. In these studies, a hypothetical design was used due to the regulatory limitations using Xpert for clinical management of TB patients. Therefore, assessing the time point on which anti-TB treatment was initiated using the Xpert test result was based on the assumption that a positive Xpert result would have resulted in immediate treatment initiation.

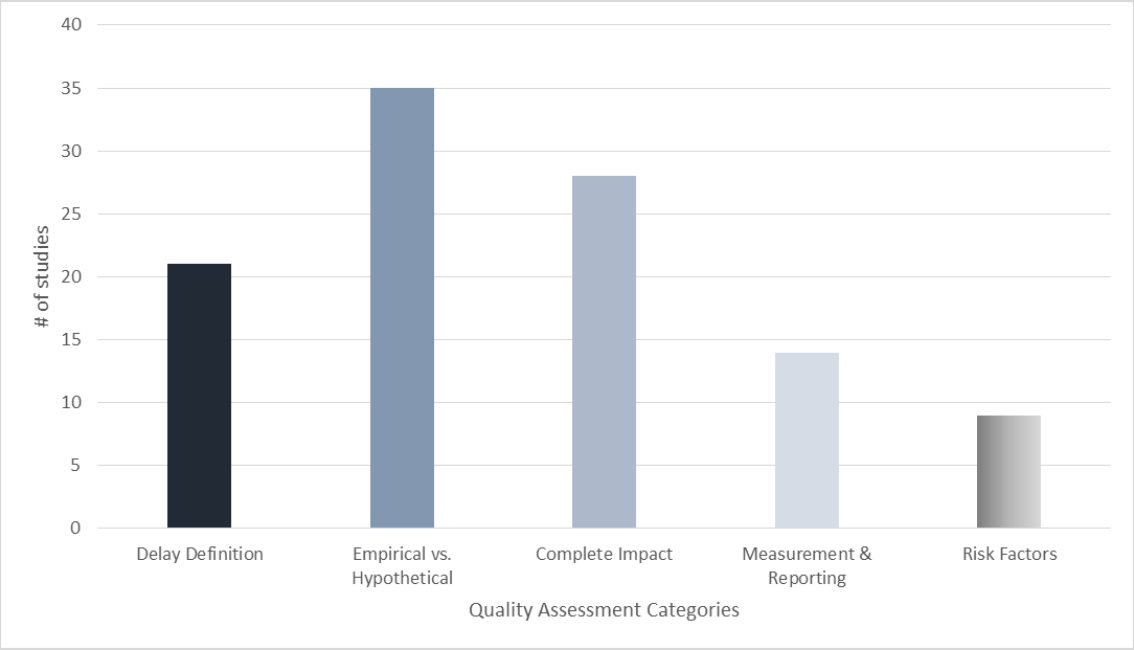


Figure 3. Overall summary of quality assessment scores
Number of studies receiving a score of '1' by quality assessment category

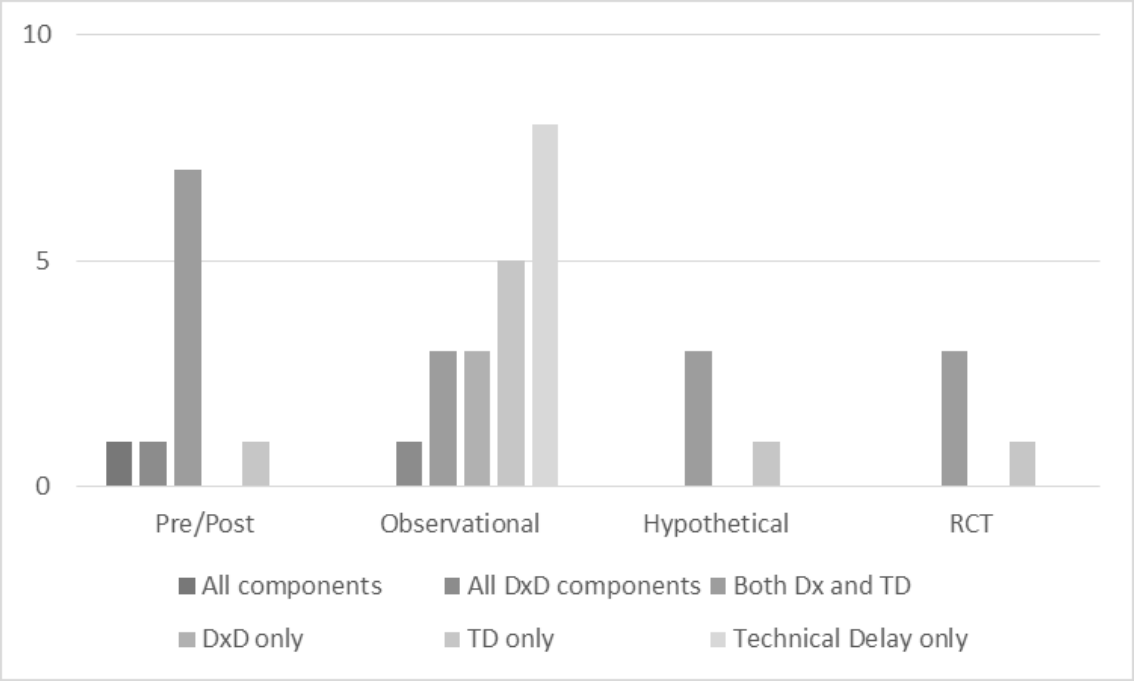


Figure 4. Types of time delay parameters reported by study design
Note: All components (Sample transport delay, Technical delay, Reporting delay, Treatment delay), All DxD components (Sample transport delay, Technical delay, Reporting delay)
* DxD = Diagnostic delay, TD = Therapeutic delay, RCT = Randomized Control Study

Impact of NAATs on diagnostic delay

Of the 39 studies included in this review, a total of 23 (59%) studies reported median time delays for both NAATs (Xpert or LPA) and comparator tests (smear, culture, or culture DST) with only 13 (33%) studies reporting p-values. Only a smaller subset of studies (9 for Xpert and 10 for LPA for meta-analysis) reported medians and IQRs for both the NAAT and a conventional test (Xpert vs. smear or LPA vs. Culture DST) with which the mean and variance could be estimated to calculate raw mean differences. The overall median diagnostic delay (IQR) for Xpert, smear, and culture for DS-TB were 1.02 days (1.00 – 5.00), 2.00 days (1.15 – 5.33), and 41.25 days (34.25 – 47.50), respectively. The delay observed between Xpert and smear were not statistically significant (p-value: 0.208). For drug susceptibility testing, overall median delay for LPA was 12.70 days (7.50 – 26.00), but was as long as a median of 29.30 days (24.90 – 37.38).

A random effects meta-analysis of mean differences showed that use of Xpert reduced diagnostic delays by an average of 2.38 days (95% confidence interval (CI): -0.09- 4.85) compared to smear (Figure 5A). When compared to culture (liquid), Xpert reduced diagnostic delays by an average of 35.58 days (95% CI 18.11, 55.09). For LPA, mean reduction in diagnostic delay was 46.57 days (95% CI 28.89, 64.23) for all studies reporting at least any one – overall estimate, direct sputum, or culture isolate – comparative time delay (Figure 5B). Sub-group analyses showed this effect estimate ranged between 23.93 days (95% CI 15.04, 34.81) and 66.58 days (95% CI 41.64, 91.52), depending on sample type used (direct sputum vs. culture isolates) and by restricting the analysis to only those studies reporting full diagnostic delay. Since diagnostic delays reported for LPA using culture isolates as the initial sample include delays associated with obtaining primary culture results (MTB detection), the greatest reduction in diagnostic delay is achieved when direct sputum samples are used

(maximum of 66.58 days, 95% CI: 41.64-91.52). Complete results from sub-group analyses are reported in Appendix B.

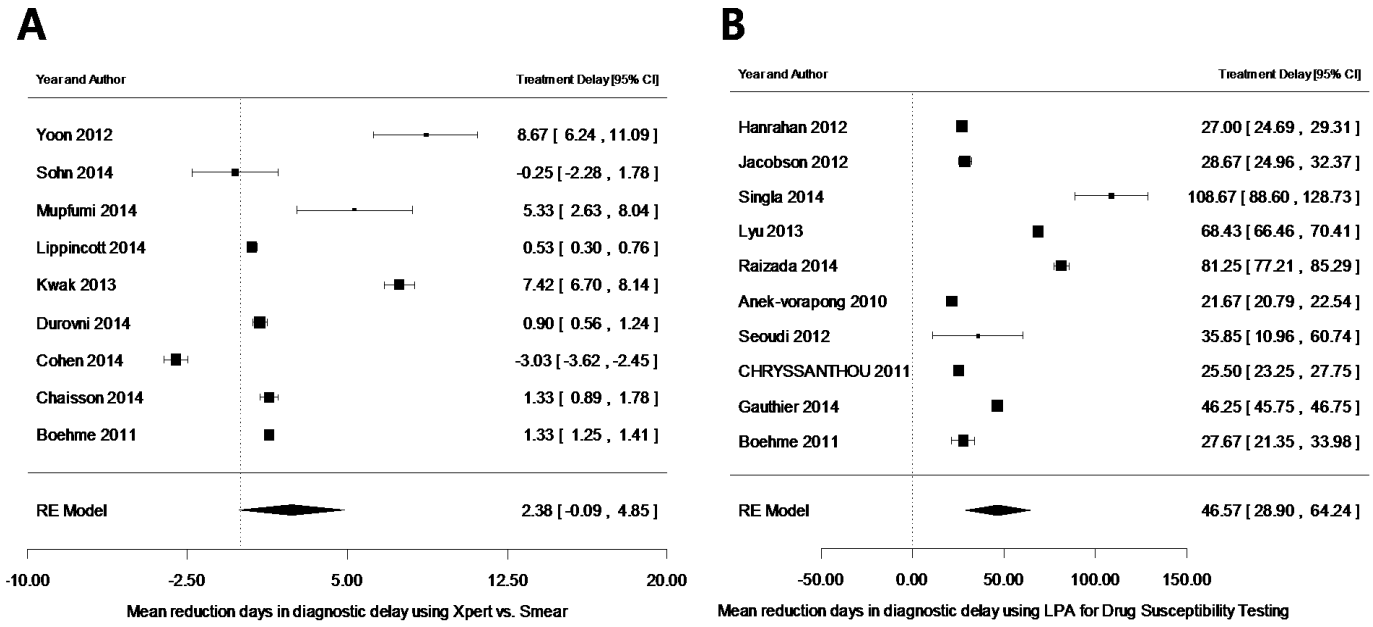


Figure 5. Forest plots of mean difference in diagnostic delay for Xpert and LPA Impact of NAATs on reducing therapeutic delays

A. Xpert vs. Sputum smear microscopy; B. LPA vs. Culture DST

A total of 21 (54%) studies reported therapeutic delays with medians and IQRs or actual median (IQR) or mean differences (95% CI) in time delay when NAATs were used for clinical management of TB patients. Two of these studies (118, 119) compared Xpert for treatment of MDR-TB patients and directly compared delay estimates with LPA. For treatment of DS-TB, overall median therapeutic delay was 2.73 days (IQR 0.09 – 5.75), while conventional diagnostic tests (smear, smear + culture, or smear + clinical diagnosis) took 11.40 days (IQR 4.92 – 28.00), which was statistically significant (p-value < 0.001). For DR-TB treatment, use of LPA could reduce median therapeutic delay to as low as 19.80 days (IQR 18.20 – 38.00) when direct sputum sample was used as compared to liquid culture,

which took a median of 80.00 days (IQR 77.00 – 86.40). When culture isolates were used for LPA, therapeutic delay was 50.00 days (IQR 41.30 – 61.80), which was not statistically significant compared to culture DST (p-value 0.052).

The pooled mean reduction in therapeutic delay for DS-TB when using Xpert (Figure 6A) as a primary diagnostic test (n = 13) was 16.54 days (95% CI 6.73, 26.35). In sub-group analyses, studies employing hypothetical design (n = 4) showed the largest mean delay reduction at 21.89 days (95% CI 5.39, 38.40). When excluding hypothetical studies and those studies that did not report overall (combined smear positive and negative TB patients) therapeutic delays (n = 6), Xpert only effectively reduced therapeutic delay by 4.75 days (95% CI 0.94, 8.57) (Figure 6B). When Xpert was primarily used for DST (n = 2), DR-TB treatment was initiated 20.44 days earlier (95% CI 10.31, 30.57) compared to LPA.

Use of LPA for DR-TB treatment (all studies reporting any one of the three therapeutic delays) also reduced therapeutic delays considerably at 62.48 days (95% CI 27.72, 97.52) (Figure 6C). Maximum reduction in therapeutic delay was achieved when direct sputum samples were used as the sample for LPA testing at 73.66 days (29.30, 118.01). When culture isolates were used as a primary specimen for LPA (Figure 6D), therapeutic delay was reduced by 43.59 days (95% CI -9.77, 96.96), but it was not statistically significant, which was consistent to our findings above. When further restricting our analysis to studies that report overall (combined estimate of direct sputum and culture isolate) LPA therapeutic delay against culture DST (n = 2), the pooled mean reduction in delay was 21.96 days (95% CI 14.54, 29.37). Of all of our meta-analyses, this particular analysis was the only analysis that showed an I^2 value indicating low heterogeneity (< 30%). All other analyses showed I^2 values greater than 85% (mostly <95%), suggesting considerable heterogeneity between the studies included in our analyses. A summary table of our meta-analysis results are listed in Table 3.

Diagnostic delay

Xpert				
Studies included	**Mean reduction in time delay days (95% CI)**	**I^2**	**Q test (p-value)**	**# of incl. studies**
Any (vs. smear)	2.38 (-0.09, 4.85)	99.77%	0.0001	9
Any (vs. culture)	36.59 (18.11, 55.09)	99.91%	0.0001	3
LPA				
Studies included	**Mean reduction in time delay days (95% CI)**	**I^2**	**Q test (p-value)**	**# of incl. studies**
Any	46.57 (28.89, 64.23)	99.91%	0.0001	10
Direct sputum sample	53.10 (33.12, 73.08)	99.93%	0.0001	8
Culture isolate sample	23.93 (15.04, 34.81)	98.61%	0.0001	5
Overall (combined estimates all sample types)	27.47 (25.50, 29.43)	0.00%	0.4547	2

Therapeutic Delay

Xpert (DS-TB)				
Studies included	**Mean reduction in time delay days (95% CI)**	**I^2**	**Q test (p-value)**	**# of incl. studies**
Any (all studies reporting therapeutic delay)	16.54 (6.73, 26.35)	99.97%	0.0001	13
Empirical studies only	4.75 (0.94, 8.57)	99.61%	0.0001	6
Xpert (DR-TB) vs. LPA				
Studies included	**Mean reduction in time delay days (95% CI)**	**I^2**	**Q test (p-value)**	**# of incl. studies**
Xpert for DR-TB vs. LPA	20.44 (10.31, 30.57)	90.57%	0.0011	2
LPA (vs. Culture DST)				
Studies included	**Mean reduction in time delay days (95% CI)**	**I^2**	**Q test (p-value)**	**# of incl. studies**
Any	62.48 (27.72, 97.24)	99.11%	0.0001	7
Direct sputum sample	73.66 (29.30, 118.01)	99.06%	0.0001	5
Culture isolate sample	43.59 (-9.77, 96.96)	99.01%	0.0001	3

Table 3. Summary of random effects meta-analysis of reduction in diagnostic and treatment delays using NAATs for DS and DR-TB

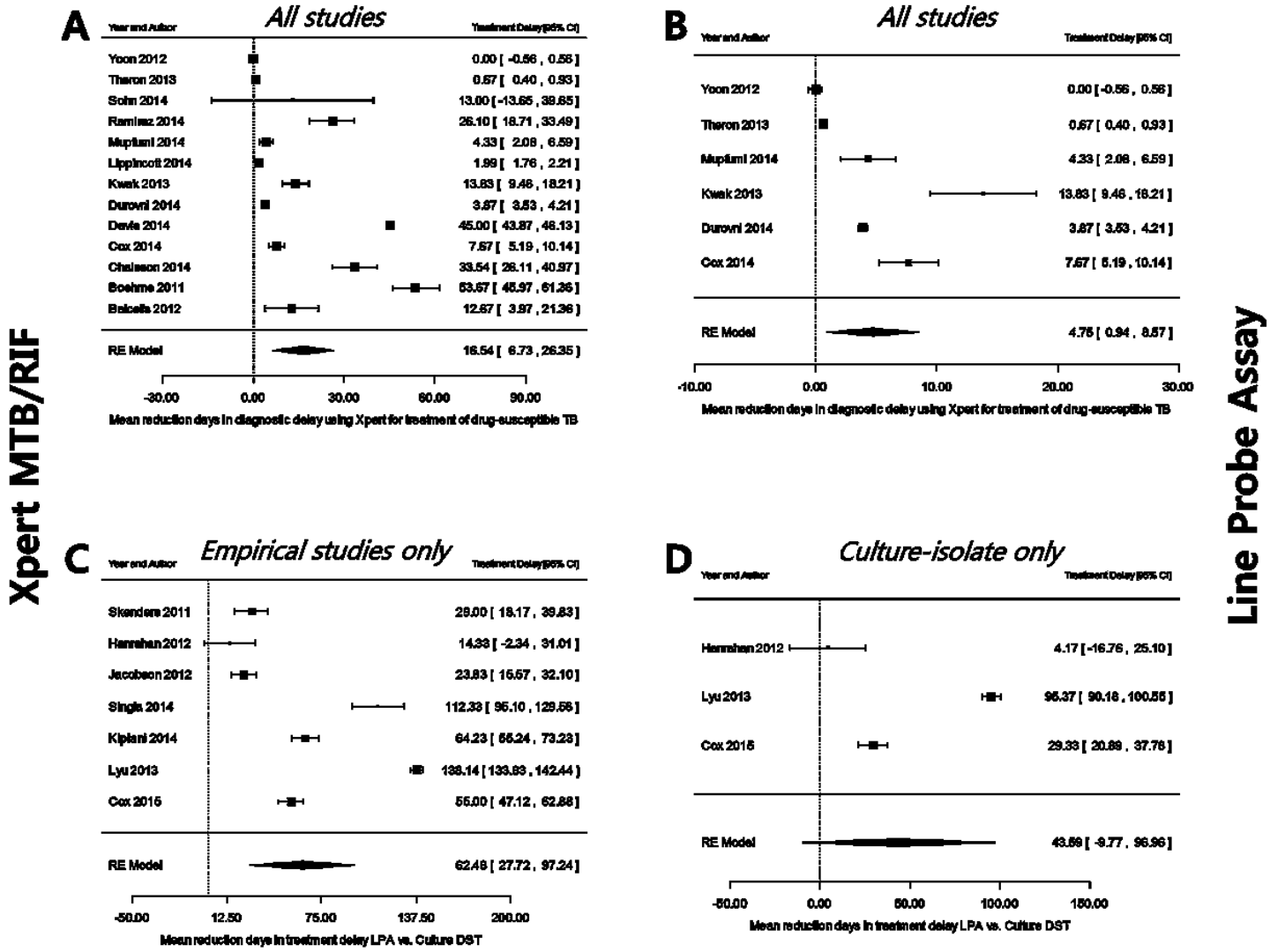


Figure 6. Forest plots of raw mean difference in therapeutic delay for Xpert and LPA

A. Xpert, including all studies (if not reporting overall figure, include lowest delay estimate), B. Xpert, only empirical studies reporting overall (combined smear positive and negative TB) diagnostic delay , C. LPA, including all studies (if not reporting overall figure, include lowest delay estimate), D. LPA, only studies reporting culture isolate for LPA testing

Discussion

One of the advantages of TB-NAATs (LPA and Xpert) is that these technological platforms are capable of accurate bacteriologic diagnosis of TB within the "same day" of sample processing (44, 136). This is in contrast to conventional methods of diagnosing TB that often delay proper diagnosis and treatment until months after a patient's first clinical evaluation for TB. Thus, it can be hypothesized that the routine use of NAATs should translate to significant reductions in time delays to diagnosis and treatment. Our meta-analysis of mean delay reductions suggests that use of TB-NAATs reduced both diagnostic and therapeutic delays for patients investigated for DS or DR forms of TB, although there was considerable heterogeneity among studies.

Of the Xpert studies reporting complete empiric therapeutic delay (excluding hypothetical studies), all but one study (113) reported a time delay reduction of less than 14 days. Seven of these studies reported insignificant or less than 7 days of time difference in therapeutic delays when comparing Xpert interventions to the comparator strategy (generally AFB smear, clinical diagnosis, and/or culture result) (66, 101-104, 106, 117). All of the studies reporting therapeutic delays for Xpert were conducted in highly restrictive and controlled clinical trials of which 5 were RCTs. In these studies, high rates of empirical treatment (> 50%) and diagnosis (generally a combination of clinical symptom evaluation, CXR, and anti-biotic trials) (137) were also reported. This has several important implications in interpreting Xpert's impact on time delays in diagnosing and treating TB.

First, in settings of high empirical diagnosis and treatment, it is likely that more non-TB patients will be prescribed TB treatment at a much earlier stage. This seemingly improves the performance of the baseline practices while representing, in part, over-treatment. When compared with an intervention aimed to improve time delays such as using rapid diagnostic

tools (e.g. Xpert), it may bias the effect estimate towards the null (138). This is confirmed in the multi-center RCT study where greater a proportion of patients in the smear arm were put on treatment based on empirical evidence (197/758 vs. 130/744) (66). This effectively abrogates any potential time delay benefits that the Xpert intervention could have had in those patients who had smear negative TB disease.

Second, operational and clinical management issues imply greater implications in realizing the potential benefit of the "same-day" rapid diagnostic tests as highlighted in several studies included in this review (91, 112-114, 118, 119, 121, 122). In Brazil, diagnostic delays in the Xpert arm were not reduced when compared to the smear arm, and additionally required more than 7 days for the test result to be reported for clinical use (103). This is largely due to the operational problems (health system infrastructure, lack of guidelines and algorithms for management of certain patients with discordant NAAT results, etc.) observed early in the nation-wide implementation of Xpert (139). A study by Cox and colleagues confirmed that time delays in diagnosis and treatment post-implementation of a newer technology can be improved over time with focused efforts to improve program implementation (119). In this study, therapeutic delays in DR-TB treatment were reduced by 38% over the one year roll-out of Xpert in South Africa between 2012 and 2013. A substantial post-implementation reduction in therapeutic delay was observed for LPA in the same study where it was reduced from 76 days (62-111) during the first year of limited decentralized implementation (2007) to 28 days (16-40) 3 years after the improved implementation program was initiated in 2009.

Several studies assessing key sub-components of time delays in diagnosis and treatment initiation showed specific logistical and operational issues caused significant delay for the "same-day" diagnostics. Centralized testing of Xpert at high volume laboratories created more delays (6.3 days) than smear microscopy (3.3 days), largely caused by slower sample

processing and result transfer to relevant clinics (114). The significance of operational and logistical consequences was greater in LPA as shown in one study, where more than 28 days of delays (of 38 total therapeutic delay days) were attributable to the components other than LPA laboratory processing (5 days) (124). In two studies from South Africa, a "same-day" DST test resulted in therapeutic delays of more than 50 days, even when LPA was performed directly from sputum samples (121, 122). Any reduction of time delays was nullified when culture isolates (smear negative presumptive MDR-TB patient) were tested on LPA, taking more than 73.5 days (vs. 78 days for culture DST) (121), consistent with the results of the meta-analysis including these two studies. Considering the significance of the problem of MDR-TB in South Africa, these findings are alarming and highlight the importance of evaluating operation and patient level factors that can influence delays in diagnosis and treatment of TB.

Our findings also highlight the need for methodological improvement and standardization in assessing time delays for diagnosis and treatment of TB. Our review found that in many instances, authors used the same terminologies to define different components of the time delays. For instance, "turnaround time", "time to detection", and "laboratory processing time" were used to describe the time from specimen receipt to test result (laboratory specific time components) for 11 studies (89, 114, 122, 124, 126-128, 131-134) while 10 studies (100, 106, 112, 115, 116, 118, 120, 121, 129, 135) employed these terminology to define complete diagnostic delay. Measurement starting times widely varied and were often unclear with some studies reporting "first contact with the health system" or "identification of patients suspected of MDR-TB" as the starting point of time measurement. Only 1 study reported a complete 4 components of diagnostic and therapeutic delays (124) and more than half (n = 22) of the studies reported diagnostic or therapeutic delays as a single estimate without providing sub-

components of the delay. Our findings are consistent with other systematic reviews (6, 32, 85) that suggest that inconsistencies and lack of consensus in defining components are critical obstacles in interpretation and generalizability of study findings.

Moreover, statistical analyses of time delay outcomes were inappropriately conducted or inadequately reported in the majority (64%) of the studies. Time delay estimates were reported in 'means' with confidence intervals rather than 'medians' with IQR without description of the distribution of the time delay data. Since normality assumption is required for computation of means and confidence intervals, this can be highly problematic for reporting time estimates, which typically are skewed and non-normally distributed. Some studies reporting medians did not include IQRs or reported means with IQRs. Additionally, only 15 studies (38%) reported p-values for tests of significance comparing time delay estimates between NAAT and conventional tests. Only 9 studies (23%) reported and assessed (analytically and by design) risk factors for time delays, further limiting our understanding of the true impact of NAATs on time delays in diagnosis and treatment of TB. These methodological limitations observed in our review may be due to the fact that for most studies, time delay outcomes were second and tertiary outcomes, and thus proper analytical assessment of these outcome measures may have been overlooked.

Our review is the first systematic review to summarize the comparative impact of NAATs on time delays in diagnosis and treatment of DS and DR-TB using a random effect meta-analysis for absolute mean delay differences. Our meta-analysis results are consistent with the findings observed in the narrative assessment. However, our meta-analysis results should be interpreted with caution. First, conversion of median delay estimates into means required an assumption of normal distribution where time delay estimates are non-normally distributed and positively skewed. This causes uncertainty in the effect estimate to be widened, where in

some instances, a ‘negative’ lower bound CI was reported for some delay measures. However, it is difficult to assess the direction or magnitude of the normality assumption that our conversion of medians into mean would have on the overall meta-analysis results. Second, considerable statistical power was lost when restricting studies that presented necessary data for mean conversion and direct comparison of delay parameters. Third, high level of heterogeneity across all of our meta-analysis suggest considerable limitations in generalizability of our pooled estimates.

It has long been understood that the delays in diagnosis and treatment of TB are mainly caused by the vicious cycle of repeated healthcare facilities providing inadequate care for TB (6, 32). This problem is an inevitable consequence of the prolonged operational and technical challenges TB control programs have faced in the LMICs. On a positive note, the global roll-out of Xpert dramatically changed the landscape of TB diagnosis in many LMIC and has led to considerable improvements in TB diagnostic infrastructure, quality of the TB control programs, and stimulated the development of novel tests that could reach much closer to the patient than Xpert (140). These positive trends will potentially have positive impact on improving time delays downstream of diagnosis and the burden of the TB disease. However, the magnitude of such effects will depend heavily on the characteristics of health system operations, as well as patient and provider behaviors in each setting (141). Therefore, there is an urgent need to assess the effect and mechanisms of societal and health system factors influencing overall delays in TB care post implementation of NAATs; this would facilitate the creation of better-informed policy and future implementation strategies, and maximize the benefit of existing and future rapid diagnostic tools.

Chapter 4: Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact

Sohn H et al. *Clin Infect Dis* 2014

4.1 Preface

At the time of our study, Xpert MTB/RIF assay (Xpert) had just been approved by WHO and there were only a limited number of studies assessing the performance of the Xpert MTB/RIF (44, 142-144). While there was much hype and excitement about the test, Xpert needed to be further evaluated for several reasons: 1) Although the technology is WHO-approved, there are virtually no data on how the test performs in low TB incidence settings (e.g. Canada), where most TB cases are smear-negative, and where relatively high prevalence of non-tuberculous mycobacteria (NTM) creates challenges for smear microscopy; 2) Although the technology has the potential to be used at the point-of-care, currently there are no published data on its feasibility and potential clinical impact when the technology is placed in a clinic setting, and used by routine clinic staff with minimal training; 3) Because cultures take 2 - 3 weeks, the Xpert test has the potential to reduce diagnostic delays in smear-negative TB cases, but no published data existed on potential reduction in diagnostic delays that can be achieved using the Xpert assay.

Thus, in this manuscript, we assessed diagnostic performance and feasibility of a point-of-care (POC) Xpert testing strategy in a routine high throughput TB clinic in Montreal, Canada. As the use of Xpert for diagnosis and clinical management of TB patients was not approved during our study, we hypothetically assessed the potential impact of POC Xpert strategy on reducing delays in diagnosis and treatment, on the basis of the idea, ‘if Xpert MTB/RIF had been used for initiation of treatment, when would TB treatment have begun?’

4.2 Manuscript

Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact

Clin Infect Dis. 2014 Apr;58(7):970-6.

Hojoon Sohn MPH[1,2], Abebech D. Aero PhD[1,2], Dick Menzies, MD MSc[1,2], Marcel Behr, MD MSc[1,2], Kevin Schwartzman, MD, MPH[1,2], Gonzalo G. Alvarez MD, MPH[1,3], Andrei Dan[1], Fiona McIntosh[1,2], Madhukar Pai*, MD PhD[1,2], Claudia M. Denkinger*, MD PhD[1,4]

Affiliations

[1] McGill International TB Centre & Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

[2] Respiratory Epidemiology & Clinical Research Unit, Montreal Chest Institute, Montreal, Montreal, Canada

[3] Ottawa Hospital Research Institute, University of Ottawa, Ottawa, Canada

[4] Division of Infectious Disease, Beth Israel Deaconess Medical Centre, Boston, USA

Abstract

Rational: Xpert® MTB/RIF, the first automated molecular test for tuberculosis, is transforming the diagnostic landscape in low-income countries. However, little information is available on its performance in low-incidence, high-resource countries.

Methods/Objectives: We evaluated the accuracy of Xpert in a university hospital TB clinic in Montreal, Canada, for the detection of active TB on induced sputum samples, using mycobacterial cultures as the reference standard. We also assessed the potential reduction in time to diagnosis and treatment initiation.

Results: We enrolled 502 consecutive patients who presented for evaluation of possible active TB (most with abnormal chest radiographs, only 18% symptomatic). Twenty-five subjects were identified to have active TB by culture. Xpert had a sensitivity of 46% (95% confidence interval (CI): 26-67) and specificity of 100% (CI: 99-100) for detection of *Mycobacterium tuberculosis*. Sensitivity was 86% (CI: 42-100) in the 7 smear-positive subjects, and 28% (CI: 10-56) in the remaining smear-negative, culture-positive subjects; in this latter group Xpert results were obtained a median 12 days before culture results. Subjects with positive cultures but negative Xpert results had minimal disease: 11 of 13 had no symptoms on presentation, and the mean time to positive liquid culture results was 28 days (CI: 25-47 days; compared to 14 days, CI: 8-21days in Xpert/culture-positive cases).

Conclusion: Our findings suggest limited potential impact of Xpert testing in high-resource, low-incidence settings due to lower sensitivity in the context of less extensive disease, and limited potential to expedite diagnosis beyond what is already achieved with the existing, well-performing diagnostic algorithm.

Introduction

The Xpert® MTB/RIF assay ("Xpert"; Cepheid, Sunnyvale, USA) is an automated nucleic-acid-amplification test for sputum specimens that can detect both *Mycobacterium tuberculosis* (MTB) and rifampin resistance within two hours, and requires minimal hands-on time. When tested in high-incidence settings, usually with spontaneously expectorated sputum, Xpert is highly accurate, (sensitivity of 88% and a specificity of 98%) (145). Due to its excellent performance characteristics, Xpert is transforming the diagnostic landscape in the developing world and is now being used in over 80 countries (45).

Xpert has also been recently approved by the US Federal Drug Administration (FDA) and Health Canada (146). Nevertheless, it is conceivable that important factors in evaluating the performance characteristics of the test such as patient population, stage of disease, methods for obtaining sputum sample (spontaneous vs. induced) and accuracy of routine smear and culture tests may differ between high- and low-resource settings. Yet, the performance of the test has not been studied in routine use in high-resource, tertiary-care settings with low-incidence of TB (147, 148). It is therefore critical to generate evidence on whether existing data and policies are transferrable to these settings (45).

In Canada, the current TB incidence is 4.6 per 100,000 population, with most cases among immigrants (3). Most pulmonary TB disease in Canada (as in other low-incidence settings) is smear-negative (66%), and therefore diagnosed only by liquid culture-based techniques that typically take 2-3 weeks to provide a result (149). Delays in diagnosis and treatment can increase patient morbidity and mortality (6, 32). While smear-negative cases are less infectious than smear-positive, they may account for up to one fifth of all secondary transmission (36, 150). Furthermore, the suspicion of TB has economic implications for the health care system, as patients may be hospitalized for respiratory isolation, while undergoing the relevant investigations.

The Xpert assay may enhance accurate and rapid detection as it can detect up to 67% of smear-negative cases (145). In addition, it might be suitable for use at the point-of-care as the test's sample reagent has potent tuberculocidal properties, thus largely eliminating biosafety concerns (43). With the use of Xpert at the point-of-care and the availability of results within hours, patients can potentially be diagnosed with TB at their first visit, which would

conceivably shorten the time to treatment and reduce transmission. However, there are limited data on the use of Xpert at the point-of-care, outside of laboratories (110, 151).

With this study, we aim to improve the understanding of the accuracy and the potential impact of Xpert in a low-incidence, high-resource setting.

Methods

Study participants

Consecutive patients aged ≥18 years, referred to the Montreal Chest Institute TB Clinic for evaluation of suspected active pulmonary TB were recruited. The institutional review board of the McGill University Health Centre, Montreal, Canada, approved the study.

Specimen collection/processing

Our TB Clinic policy is to collect induced sputum samples, using 3% hypertonic saline solution and an ultrasonic nebulizer on all patients with possible/suspected active TB. Two samples were obtained from all consenting patients on the day of enrollment. The first sample was processed in standard fashion in the clinical microbiology laboratory, including smear (Auramine O method) and liquid culture (Mycobacterium Growth Indicator Tube, Becton Dickinson, USA). The second sample was used for Xpert testing.

The Xpert test was performed at the TB clinic according to the standard protocol for unprocessed samples, per the manufacturer (152). Further information is available in Appendix C and Figure E1.

When we started the study, Xpert had been endorsed by the World Health Organization (153). Approval of the test by Health Canada followed in 2012. Since Xpert was done outside the hospital-approved clinical lab, test results obtained as part of the study were not made available for clinical decision-making. However, following approval of the test by Health Canada (after enrollment of 394 subjects), the microbiology laboratory was alerted of any positive Xpert result and a conventional nucleic acid amplification test (NAAT; Cobas TaqMan MTB, Roche Diagnostics, Switzerland) was suggested (if not already done).

The reference standard was liquid culture on three processed samples, followed by phenotypic culture-based drug-susceptibility testing (DST) at the provincial reference laboratory (154). For all discrepant results (i.e. rifampin resistant on Xpert but sensitive on DST) sequencing of the rpoB gene was performed (Appendix D). From January 2012 all MTB isolates were also routinely typed using mycobacterial interspersed repetitive units (MIRU) typing. From this data, we determined that one positive culture result (Xpert-negative) could have been due to cross-contamination in the laboratory and therefore was excluded from all analyses.

To assess the limit of detection of Xpert and the potential impact of hypertonic saline on the performance of Xpert, we added Bacille Calmette-Guerin (BCG) at concentrations of 250 (n=6), 125 (n=8), 62 (n=8), 31 (n=10) CFU/ml to normal and 3% hypertonic saline, then submitted these samples for Xpert sample preparation and testing as above.

Statistical analysis

We calculated sensitivity and specificity of the Xpert compared to the culture reference standard. We assessed the accuracy of rifampin resistance testing on Xpert compared with culture-based DST.

We assessed clinical impact of all diagnostic methods by examining the interval from procuring the first sample to obtaining the relevant diagnostic result. Furthermore, we compared the time from first sputum collection to treatment initiation and the days of empiric treatment given prior to culture-confirmation. We compared this with the time when the Xpert result would have been available to the physician, if results had been shared (Figure E2, Appendix D).

We used Standards for the Reporting Diagnostic Accuracy (STARD) (155, 156) for reporting the study results.

Results

Between October 2011 and May 2013 we enrolled 502 consecutive patients who presented to the TB clinic for evaluation of possible active TB. The median age of subjects was 44 years (interquartile range [IQR]: 31-61) and 44% were female (Table 1). Persons referred for immigration-related screening constituted the largest number of subjects (294, 59%) and only

93 subjects were born in Canada (18.5%). Many (44%) were born in countries with very high TB prevalence (>100/100,000). Only 12 subjects were known to be HIV positive and 15 subjects had other immunocompromising co-morbidities (5%). A history of prior active TB was reported by 111 subjects (22%).

A sizeable fraction of subjects were referred for evaluation in the context of an immigration screen that yielded a chest X-ray with a possibly TB related abnormality (22%; Table 2). Others were contacts of active TB cases, who had positive tuberculin skin tests (5%). Only 18% had symptoms suggestive of active TB (i.e. fever, cough, night sweats, weight loss) and overall 74% had an abnormal chest X-ray (19% of these with findings highly suggestive of active TB, i.e. cavitation, and/or apical fibronodular disease]; Table 2).

Twenty-five subjects were identified to have active culture-confirmed TB. Eleven subjects were smear positive but only 7 of these were identified to have MTB disease (3 others had non-tuberculous mycobacteria and one was a false-positive as culture/NAAT negative).

Variables	n	%
Subjects total	502	100
Age categorized		
18-29	76	15.2
30-49	223	44.4
>50	203	40.4
Gender		
Female	223	44.4
Male	279	55.6
TB prevalence in country of birth (all forms, per 100,000)		
Canada	93	18.5
Low (<=25)	52	10.4
Medium (26-50)	14	2.8
High (51-100)	120	23.9
Very high (>100)	223	44.4
Status in Canada		
Canadian-born Citizen	93	18.5
Foreign-born Citizen	62	12.4
Immigrant	294	58.6
Foreign-born Student	12	2.4
Work Permit	12	2.4
Other	29	5.8
Co-morbidities		
Diabetes	6	1.2
Malnutrition	0	0
End-stage renal disease	1	0.2

History of malignancy	2	0.4
Treatment with immunosuppressive medications	6	1.2
HIV testing result		
Negative	37	7.4
Positive	12	2.4
History of Tuberculosis		
Active	111	22.1
Latent	9	1.8
Close contact with TB patient	31	6.2

Table 1: Demographic and clinical characteristics

Variables	**n**	**%**
Subjects total*	502	100
Symptoms		
Fever	17	3.4
Cough	85	16.0
Hemoptysis	14	2.8
Chest Pain	10	2.0
Shortness of breath	6	1.2
Night-Sweats	12	2.4
Weight-loss	23	4.6
Radiographic findings		
Apical fibronodular disease	64	12.8
Cavitation	10	2.0
Granuloma	36	7.2
Costophrenic angle blunting	14	2.8
Other abnormality	285	56.8

* For 10 subjects no clinical information is available

Table 2: Symptoms and radiographic findings

Non-interpretable results on Xpert

Non-interpretable results were detected in 44 (8.8%) samples overall. For most of these tests (37/44; 84%) the internal control failed. If Xpert yielded an invalid result, repeat testing was performed either on the same sample (if sufficient volume) or a repeat sample. All repeat testing resulted in an interpretable result (5 subjects did not return for repeat testing). While non-interpretable results decreased somewhat with the change from the G3 to the G4 cartridge (11.9% for G3, 95% confidence interval (CI): 7.1-18.4 compared to 7.5% for G4, CI 5.0-10.8%), the number of invalid results still exceeded that reported in the literature (Table

3) (3). Therefore we requested an evaluation by the manufacturer. The manufacturer discovered that one lot accounted for 91% of all invalid tests but only 70% of all tests (odds ratio 5.5 for this lot compared to all other lots, CI 1.3-23.9). Further evaluation of reasons for invalid results is ongoing with the manufacturer.

Xpert accuracy

Xpert detected 11 out of 25 subjects with culture-confirmed TB, for a sensitivity of 46% (CI: 27-67) and a specificity of 99.8% (CI: 98.7-100) for detection of culture-positive TB (Table 3). The sensitivity was improved in smear-positive subjects (86%, CI: 42-100) compared to only 29% sensitivity in smear-negative subjects (CI: 10-56). While sensitivity appeared to be lower with the G4 cartridge (33%; 12-62) compared to the G3 cartridge (67%; 30-93), the CIs were wide and overlapping (Table 3).

One subject was "false-positive" on Xpert as culture result was negative (no documented pretreatment). In that case, the Xpert result was confirmed by a positive NAAT in the clinical microbiology lab. Eight other subjects who were treated for TB based on clinical grounds (not-culture confirmed) were Xpert negative (no NAAT done).

		n	N cases[#]	Sensitivity (95% CI)	Specificity (95% CI)	Number Invalid* (%)
Xpert results (1st test)[†]		501	25	46 (26-67)	100 (98-100)	44 (8.8)
By smear result	**Positive**	11	7	86 (42-100)	100 (29-100)	1 (9.1)
	Negative	425	18	29 (10-56)	100 (99-100)	40 (8.6)
By cartridge version	**G3**	143	10	67 (30-93)	100 (97-100)	17 (11.9)
	G4	358	15	33 (12-62)	100 (98-100)	27 (7.5)

*Invalid or erroneous result; CI=confidence interval; [†]One subject with contaminated culture result excluded; [#] Culture-confirmed tuberculosis cases

Table 3: Xpert results

Rifampin resistance results on Xpert

Only 2 isolates were predicted to be rifampin resistant on Xpert testing. Culture-based DST confirmed only one of the two to be rifampin resistant. Sequencing of the rpoB gene on the isolate that provided a discrepant result between Xpert and culture-based DST, identified a mutation in the 511 locus (Pro > Leu) that is captured by probe A of the Xpert (Appendix D).

Evaluation of low sensitivity

Evaluation of the limit of detection of Xpert yielded 100% detection of BCG at a concentration as low as 62 colony forming units (CFUs)/ml and 80% at a concentration of 31 CFU/ml in normal saline, thus suggesting an even lower limit of detection than what was described in the original validation studies on sputum samples (2). Furthermore, the sensitivity of the Xpert was the same in samples with hypertonic saline as with normal saline.

Most participants with culture-positive TB had minimal disease (Table 4). This is suggested by the fact that only 7 out of 25 (28%) culture-positive subjects were smear-positive, only 12 (44%) had symptoms at presentation and 2 subjects had no radiographic abnormalities at all. Two out of seven subjects (18%) who had only one positive culture (out of 3) were Xpert positive, while 9 out of 12 subjects (75%) with three positive cultures were Xpert positive. In addition, a longer period to culture positivity was noted for Xpert-negative, culture-positive subjects (28 days, CI: 25-47 days compared to 14 days, CI: 8-21days in Xpert-positive/culture-positive cases), suggesting a lower bacillary load. The mean cycle-threshold (CT)-value for all Xpert- and culture-positive subjects also was high at 28.2 (standard deviation of 2.9) suggesting a low bacillary burden even in those subjects who were Xpert positive (157). The presence of symptoms upon enrollment into our study was the one variable that was predictive of Xpert positivity (Table 4).

Characteristics		Total TB cases	Xpert positive	
		n*	% (n)	95% CI
Age	<35	13	46 (6)	17-77
	>35	11	46 (5)	19-75
Gender	Female	6	33 (2)	4-78
	Male	18	50 (9)	26-74
Country of origin	Canada	0	0	0
	Other	24	46 (11)	26-67
Prevalence in country of origin	Low/medium	7	43 (3)	10-82
	High/very high	17	47 (8)	23-72
History of TB	No	22	46 (10)	24-68
	Yes	2	50 (1)	1-98
Immunocompromising illness	No	24	46 (11)	26-67
	Yes	0	0	0
Symptoms	**No**	**13**	**15 (2)**	**2-45**
	Yes	**11**	**82 (9)**	**48-98**
Radiographic abnormalities	No	2	0	0-84
	Yes	22	50 (11)	28-72
Number of cultures positive	1 positive	7	29 (2)	4-71
	2-3 positive	17	53 (9)	28-77
Time to liquid culture positivity	>3 weeks	13	23 (3)	5-54
	<3 weeks	11	73 (8)	39-94

*One subject with positive culture and invalid Xpert result.

Table 4: Xpert result by subject characteristic
The only variable for which confidence intervals do not overlap is marked in bold letters.

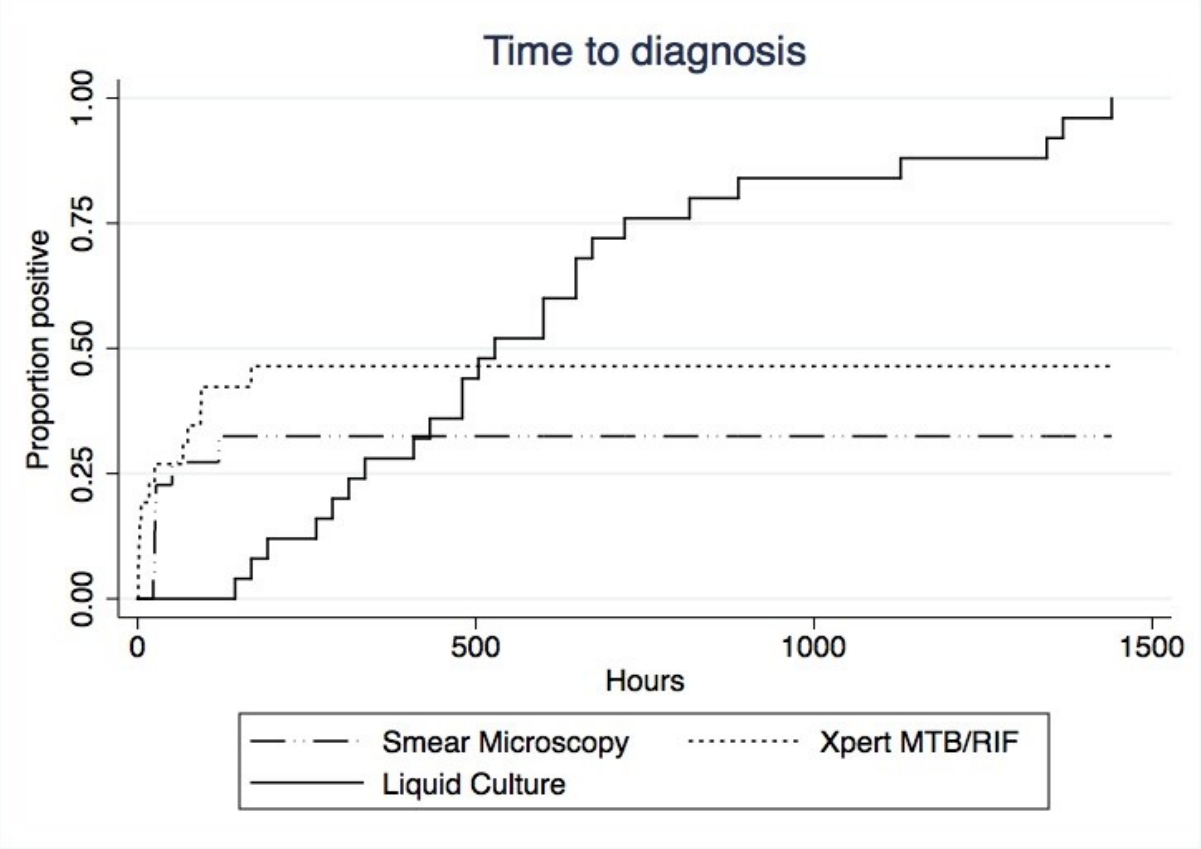


Figure 1: Sensitivity and time to positivity of different diagnostic methods
Abbreviations: MTB, Mycobacterium tuberculosis, RIF, rifampin.

Potential clinical impact evaluation

The Xpert result for all subjects was available on average within two hours. However, given that subjects were not always enrolled on initial presentation to the TB clinic, the time between the first sample and the positive Xpert result for culture-confirmed cases was a median of 25 hours, (IQR 3-93; N = 11). A positive smear result was reported within 26 hours (IQR 25-51), while a positive culture result was reported after 516 hours (i.e. 22 days; median; IQR 336-720 hours).

The actual time to treatment initiation (from initial sample provided) was 1 day for smear-positive cases (median; IQR 0-2) and 26 days for smear-negative cases (median; IQR 4-30). For 13 of the 18 smear-negative cases, Xpert was negative and therefore would have not influenced treatment decisions. For the remaining five subjects who were smear-negative but Xpert positive, treatment would potentially have been started a median of 12 days (IQR 4-23)

sooner, if results had been shared with the physicians. Treatment initiation would have been only one day earlier, at best, for smear-positive cases.

Subjects with smear-positive results who were ultimately identified not to have TB in this study were not started on TB therapy while awaiting culture results, likely because species confirmation by existing NAAT in the clinical lab was usually done within a day of the positive smear, and suspicion of clinicians was low. Thus Xpert would not have had any impact in preventing unnecessary TB treatment and possibly contact investigations in these subjects.

Discussion

The Xpert assay has been shown to effectively and rapidly diagnose TB in low-resource settings where diagnosis has hitherto depended primarily on smear microscopy, thereby potentially decreasing morbidity associated with diagnostic delay, dropout and mistreatment even if some persons with smear-negative active TB are still missed because of imperfect sensitivity . However, the impact of the technology in low-incidence, high-resource settings with full mycobacterial culture and drug susceptibility testing capability has not been adequately studied. With the recent FDA approval of this technology, it is important to generate evidence in low TB incidence settings.

Our study highlights that in a high-resource, tertiary-care setting, subjects are likely to present early in their disease course with minimal disease, in parts detected as a result of active immigration screening. This is suggested by the substantial number of asymptomatic subjects in our study and the limited number of subjects with radiographic findings consistent with active TB. Furthermore, the time-to-positivity of mycobacterial cultures in our participants was longer than the expected average for liquid cultures, suggesting a low bacillary load—notably in those with negative Xpert results (158).

The preponderance of paucibacillary disease likely accounts for the limited sensitivity of Xpert we observed. While the results for smear-positive samples are within the range of previous observations in low resource settings, the sensitivity of Xpert for smear-negative samples is substantially lower than that reported in a recent systematic review (68% sensitivity) (145). However, that review involved only subjects who were symptomatic on presentation while in our study only 18% of subjects were symptomatic.

Most studies thus far published on Xpert also used expectorated sputum, while all sputum samples in this study were obtained by induction. It is conceivable that the dilution of the sample in the process of sputum induction, results in even smaller numbers of CFUs in the cartridge. This may also contribute to the lower sensitivity of Xpert in our setting, particularly since all samples sent to the clinical microbiology laboratory for smear microscopy and culture undergo a concentration step, but for Xpert performed in the clinic, no concentration step was used (with the intent to minimize processing steps and equipment as well as biosafety concerns). An effect of hypertonic saline (used for the sputum induction) on the performance of Xpert appears unlikely as the pH of the sample obtained with induction is likely to be only minimally different from expectorated sputum. Furthermore, an evaluation of 32 samples using BCG to compare normal and hypertonic saline did not show any difference.

A decreased sensitivity of Xpert in induced sputa has also been described in preliminary results from a study of South African adults (152). However, studies in children have shown adequate sensitivity of Xpert in induced sputum(159). It is conceivable that adults with paucibacillary disease are more likely not to produce sputum, while children may have many more reasons why they cannot provide a spontaneous sputum sample (e.g. inability to follow instructions), which could explain the discrepant finding.

Concerns have been raised about the limited specificity of Xpert for rifampin resistance detection and thus its positive predictive value in a setting with low prevalence of multi-drug resistance (160, 161). In this study, only two subjects were labeled as having rifampin resistant TB by Xpert, of which only one had confirmed resistance on culture-based DST. Sequencing of the rpoB gene of the isolate with the discrepant result suggested a mutation that was associated with increased failure and relapse rates in recent studies (162, 163). This finding raises some concern about the predictive validity of phenotypic testing for rifampin susceptibility, and its use as the gold standard for confirmation of the Xpert rifampin resistance test. Sequence for confirmation of rifampin resistance detected on Xpert is therefore recommended (164).

In addition, our study highlights the limited potential impact of Xpert on time to diagnosis and likely also on treatment decisions in a setting where 1) the standard diagnostic algorithm with smear and culture, supplemented by confirmatory NAAT in the laboratory, performs

well; 2) there are excellent logistics for transport of samples and communication of results; and 3) physicians are experienced in the diagnosis and care of TB subjects (in our TB clinic, all subjects are seen by pulmonologists). However, given that Xpert was not always done on first encounter, results were not used for actual clinical decision-making and the overall number of positive cases was low, this conclusion has to be interpreted with some caution.

Furthermore, it is also unlikely that Xpert will serve as a rule-out test when cultures are available, as the overall sensitivity in smear-negative TB was low and clinicians are unlikely to withhold treatment based on a negative Xpert result if clinical suspicion remains high. However, it remains possible that Xpert may have a role in more remote areas or confined populations within a high-resource country, particularly if there is a substantial community burden of TB, and the available diagnostic infrastructure is limited. A study is currently underway to examine the role of Xpert in Aboriginal communities in the Canadian Arctic. Preliminary findings support a potential role for the new technology in this remote setting, where there is limited on-site laboratory capacity (Personal communication G. Alvarez) (3). Furthermore, Xpert may be useful in an inpatient setting, where patients typically present later in their disease. In this setting, Xpert may also reduce the time in respiratory isolation for patients suspected for having TB, and thus result in cost-savings (164, 165)

In summary, we found that the impact of Xpert testing in a low-incidence, high-resource setting is limited. These findings underscore a recommendation in the recently updated Canadian TB Standards which allows the use of Xpert MTB/RIF in laboratories, but cautions that the use of Xpert should not replace conventional smears and cultures, and recommends that all Xpert results should be confirmed by routine smears and cultures (149).

Chapter 5: Cost-effectiveness of molecular diagnostic test for tuberculosis for use in settings of low drug resistant tuberculosis

5.1 Preface

Loop-mediated isothermal amplification (LAMP) is the basis of a NAAT method, which a commercial version, the Loopamp^{TM} MTBC Detection Kit (TB-LAMP), can extract and amplify *Mycobacterium tuberculosis* specific DNA less than 2 hours at 67°C. The amplified product is detected and quantified by its turbidity, visualised by inspection or under ultraviolet (UV) light (Appendix D, Figure A1). Earlier studies have shown TB-LAMP's diagnostic performance to be equivalent to Xpert when performed in LMICs with high TB burden (166). The low infrastructural requirement, equipment cost, and relative simple and contamination-free procedures make this test potentially suitable for large-scale implementation at peripheral microscopy laboratories in low resource settings where existing molecular tests, including Xpert, may not be used.

In this manuscript, we compare the cost-effectiveness of TB-LAMP for diagnosing pulmonary TB to that of existing diagnostic strategies (e.g. sputum smear microscopy, followed by clinical diagnosis). The comparison was performed by a simulated decision analytic model using a hypothetical cohort of 10,000 patients clinically suspected of having pulmonary TB in 2 medium to high TB incidence countries with low rifampin resistance (multi-drug resistance TB prevalence). We assess two diagnostic strategies for TB-LAMP assay implementation: as a replacement test or an add-on test to sputum smear microscopy.

Results from this study have been presented to the WHO as the primary health economic evidence for the WHO policy and guideline development of use of this test.

5.2 Manuscript

Cost-effectiveness of TB-LAMP assay for the diagnosis of pulmonary tuberculosis

Hojoon Sohn[1,2,3], Frank Cobelens[4,5], Anja H. van't Hoog[4,5]

Affiliations

1. Department of Epidemiology, Biostatistics, and Occupational Health, McGill University
2. McGill International TB Centre
3. Research Institute of the McGill University Health Centre
4. Amsterdam Institute for Global Health and Development
5. Department of Global Health, Academic Medical Center, University of Amsterdam.

Abstract

Background: Loop-mediated isothermal amplification test for tuberculosis (TB-LAMP) is a manual assay (Eiken Chemical Company Ltd. Tokyo, Japan) based on nucleic acid amplification technique (NAAT) for detection of *Mycobacteria tuberculosis* (MTB) in sputum. Though expensive compared to sputum smear microscopy, ease of use, low equipment cost and maintenance, and better diagnostic performance might make this assay potentially suitable for use in settings with limited laboratory infrastructure.

Methods: We used decision analytic modelling to simulate a cohort of patients diagnosed with TB in settings where Xpert MTB/RIF (Xpert) testing is not offered, parameterized for two countries with low multi-drug resistant tuberculosis (MDR-TB prevalence) with high and low TB-HIV co-infection: Malawi and Vietnam. We compared two TB-LAMP strategies (replacement for sputum smear microscopy (SSM) and an add-on test to SSM in smear negative patients) to the base case algorithm with SSM followed by clinical diagnosis in those patients with negative SSM.

Results: In both Malawi and Vietnam, both of the TB-LAMP scenarios improved case detection rates to between 74-76% and 88-90%, respectively, compared to the base-case scenario rates of 59% and 82%. The incremental cost per disability adjusted life years (DALY) for the TB-LAMP replacement for SSM strategy was between $41 and $131, which was higher than that of the add-on scenario at $39 and $123 in Malawi and Vietnam, respectively. Both strategies were cost-effective when comparing to the World Health Organization (WHO) willingness-to-pay (WTP) threshold levels. These conclusion did not change in a range of sensitivity analyses performed.

Conclusion: Our findings suggest that TB-LAMP is potentially a cost-effective alternative to the base case of SSM *plus* clinical diagnosis in settings where Xpert is not available or is highly expensive to implement. However, given TB-LAMP's lack of capacity to detect DR-TB, financial constraints in low income countries, and emergence of novel automated point-of-care molecular tests for TB, policy makers must cautiously evaluate operational and financial feasibility prior to introducing this technology.

Introduction

Despite a decline in the global tuberculosis (TB) incidence, TB remains one of the most significant infectious disease problems globally. In 2013, an estimated 9 million people developed TB, 1.1 million of whom were co-infected with the human immune-deficiency virus (HIV), with more than 1.5 million dying of the disease (2). An estimated one-third of these cases went undiagnosed or unreported, partially due to shortcomings of conventional diagnostic tests, but also poor diagnostic and health care coverage for patients suspected of TB.

Currently, sputum smear microscopy (SSM) is the most widely used method by National TB Programs (NTPs) in resource limited settings due to the low cost of the technique and the minimal laboratory infrastructure required. However, performance of this century-old test is highly dependent on the level of training and types of microscopy equipment available (167). Molecular assays based on nucleic acid amplification techniques (NAAT) such as Xpert MTB/RIF (Xpert) have been recognized as “game changing”, cost-effective solution for TB diagnosis (1). Since the WHO endorsement, unprecedented efforts have been made for its global scale-up (13), particularly in high TB burden settings. The Xpert MTB/RIF assay is an automated NAAT, and case detection and rifampicin resistance detection are both included in the cartridge. However, the current diagnostic coverage Xpert have been limited to intermediate (district) level laboratories or tertiary care settings, reducing its potential impact on delays in diagnosis and treatment of TB. Furthermore, many low-income countries are unable to afford the Xpert MTB/RIF test, and are still reliant on smear microscopy (168). A lower cost molecular assay that can performed at the microscopy center level has therefore been identified as a priority by many stakeholders (169).

Another type of novel molecular amplification method, loop-mediated isothermal amplification (LAMP) (170, 171) is a simple and highly specific, closed-tube NAAT technique that does not require complex laboratory equipment or infrastructure. A commercially available version, the Loopamp™ MTBC Detection kit (TB-LAMP) by Eiken Chemical Co., Ltd. is a contamination resistant kit (166, 170, 171), which has a promising operational feasibility in high TB burden, low resource settings, and is particularly targeted towards use in peripheral microscopy laboratories (172). Though a manual assay, TB-LAMP procedures are simple, is contamination-free and requires less than two hours to perform.

Likewise, while lacking drug resistance testing, its minimal maintenance and infrastructure requirements could make TB-LAMP a cost-favorable and performance-optimized alternative to smear microscopy in settings where affordability is an issue, where the practicality of Xpert is limited, or where drug resistant TB is not of significant concern.

Therefore, the main objective of this study was to assess the cost-effectiveness of TB-LAMP as a replacement for SSM or as an add-on test to SSM in smear negative patients, compared to the standard of care in settings where Xpert testing coverage is limited. In such settings, TB diagnosis is limited to SSM followed by clinical diagnosis for patients suspected of new infection, and culture/drug susceptibility testing (DST) for those TB patients at increased risk of MDR). We chose settings in two countries with a relatively low Multi-Drug Resistant (MDR)-TB burden – Vietnam and Malawi – to represent distinct scenarios with respect to TB-burden, HIV-TB co-infection, and cost. Our analysis took the perspective of public TB service provider.

Methods

Diagnostic scenarios

We compared two diagnostic scenarios that include TB-LAMP against a base case. In the base case, each presumptive TB patient submits two sputum specimens for examination by SSM using Zhiel Neelsen (ZN) staining method. If positive, TB treatment was started. If negative, procedures for clinical diagnosis of smear-negative TB followed, which may include chest X-ray (Vietnam) and a course of broad-spectrum antibiotics (Malawi & Vietnam), in addition to clinical judgement (Malawi & Vietnam). In all strategies, conventional culture-based (liquid culture) drug susceptibility testing (DST) was assumed for all diagnosed TB patients who had previously been treated for TB, and therefore at high risk for drug-resistance (173). In the first alternative scenario ("add-on" strategy, with TB-LAMP added to microscopy), presumptive TB patients with smear negative results had a single TB-LAMP test done, followed by clinical diagnosis if TB-LAMP is negative. In the second scenario ("replacement" strategy), all presumptive TB patients received a single TB-LAMP test on sputum instead of SSM.

Model structure and assumptions

We developed a deterministic decision-analytic simulation model (Figure 1) based on a previously described cost-effectiveness analysis model for Xpert (1). We simulated a cohort of 10,000 presumptive TB patients (patients presenting with prolonged cough with or without systemic or other symptoms suggestive of pulmonary TB) through the diagnostic pathway and TB treatment; each step was governed by specified probabilities. The simulated cohort was stratified by history of TB treatment (new versus previously treated patients), HIV status (positive versus negative), and TB drug resistance (drug susceptible versus multi-drug resistant (MDR)-TB). Costs for diagnosis and treatment were assigned to each decision pathway, and number of patients, total costs, and disability-adjusted live-years (DALYs) averted were estimated for each outcome. Analyses were done using TreeAge Pro™ 2015 Software (TreeAge Software Inc., Williamstown MA, USA).

Model parameters and assumptions

Key model parameters are shown in Table 1 and in the Appendix E. Diagnostic sensitivity and specificity of TB-LAMP were based on the systematic review presented to the World Health Organization (WHO) Expert Group for TB-LAMP (174). Based on 13 studies, the overall pooled sensitivity was 78% and spec was 98% for TB-LAMP assay. In comparison, the overall sensitivity and specificity of Xpert MTB/RIF were 89% and 99%, based on a Cochrane review of 27 studies (42). The model was parameterized for two distinct settings in low MDR-TB countries with varied Human Immunodeficiency Virus (HIV) prevalence among TB patients: Vietnam (low HIV prevalence) and Malawi (high HIV prevalence). Relevant proportions of TB patients who were 1) new or previously treated presumptive TB patients, 2) HIV positive and negative, and 3) MDR and drug-susceptible (DS) were sourced from WHO annual country reports (175). Other key parameters were adopted from the parameters used in Vassall et al, 2011 (1).

With respect to our model assumptions, we assumed that 1) the proportion of smear-negative TB cases is determined by a fixed ratio of smear-negative to smear-positive that only depends on HIV status, 2) the probability of HIV infection is independent of re-treatment status and the probability of MDR-TB is independent of HIV status, 3) persons with TB who are not diagnosed for TB at their initial effort return after 3 months and undergo the same diagnostic scenario. The rates of death, self-cure, and conversion from smear-negative to smear-positive during these 3 months were taken from the literature (Appendix E). These assumptions are applicable to all scenarios/strategies in our model.

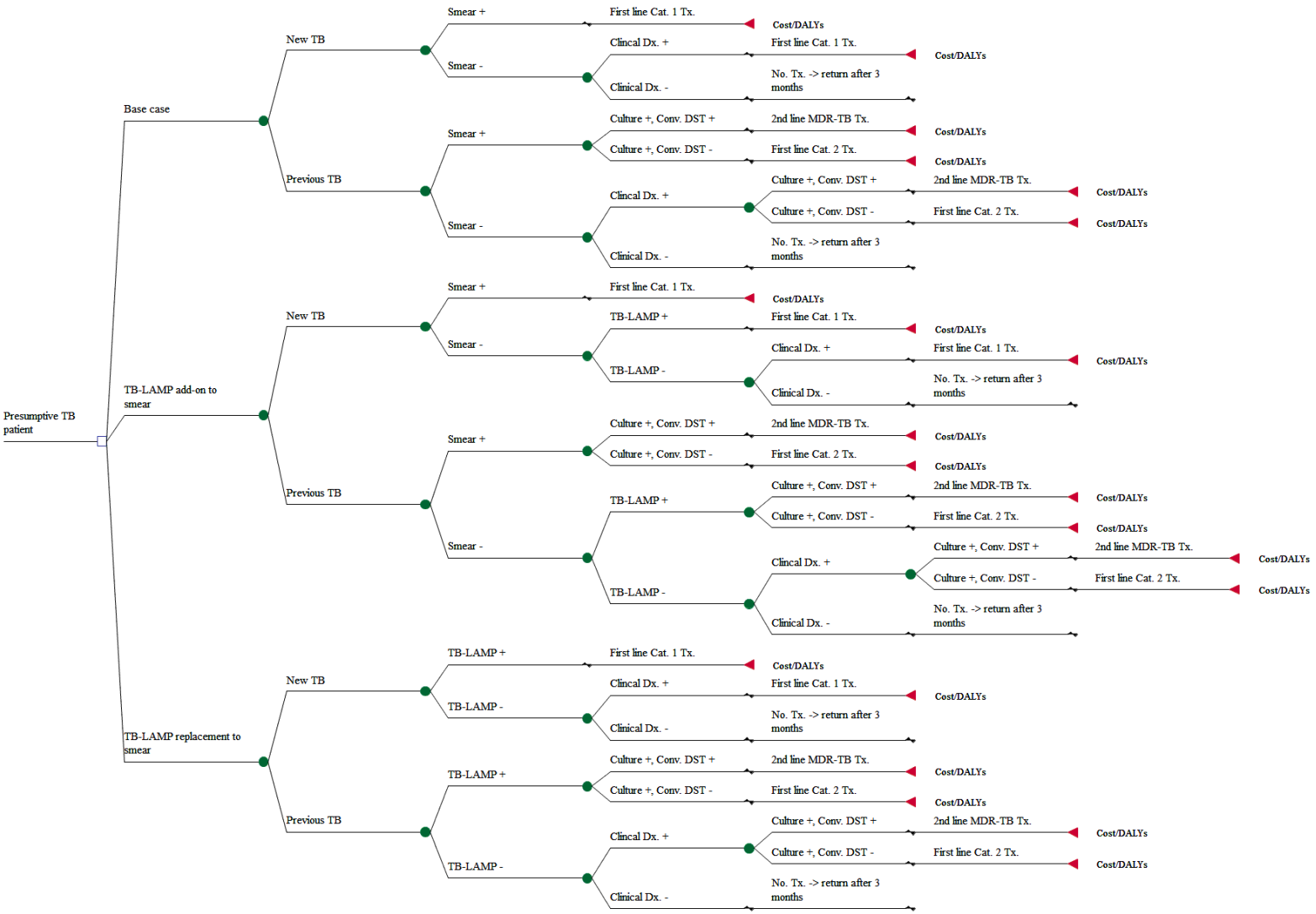


Figure 1. Simplified deterministic decision analytic model.

Concerning the diagnostic scenarios, clinical diagnosis of smear-negative TB (clinical evaluation, antibiotic trial, and/or chest X-ray (CXR) if applicable) was treated as a single diagnostic procedure with fixed test performance characteristics, as in our earlier model (1). As diagnostic performance data of clinical diagnosis were not available for Vietnam and Malawi, we referred to the data from our Xpert model (1). In that model, clinical diagnosis performance in Uganda and India was based on that the level of utilization of CXR, with India and Uganda representing low and high CXR utilization. In Vietnam, CXR is a critical component of clinical diagnosis of TB (personal communication with Dr. Nhung, NTP manager), therefore, we assumed that clinical diagnosis performance would be analogous to the Ugandan scenario in the Xpert model. Contrastingly, in Malawi, CXR is not readily available for routine evaluation of TB patients, thus, corresponding to the Indian scenario from the Xpert model (1).

Cohort proportions and diagnostic parameters	**Malawi**	**Vietnam**	**Distribution***	**Source**
Cohort proportions				
Smear-positivity 3 months after initial smear negative result	0.1 (0.015)		beta	(1)
New cases among pulmonary TB cases	0.911 (0.137)	0.878 (0.132)	beta	(2)
Multidrug resistance, among new TB cases	0.004 (0.001)	0.040 (0.015)	beta	(2)
Multidrug resistance, among previously treated TB cases	0.048 (0.018)	0.23 (0.065)	beta	(2)
HIV infection, among pulmonary TB cases	0.675 (0.101)	0.097 (0.015)	beta	(2)
Diagnostic parameters				
sensitivity for diagnosing pulmonary TB (SE)*				
TB-LAMP, smear-positive TB cases	0.95 (0.040)		beta	(174)
TB-LAMP, smear-negative TB cases	0.42 (0.075)		beta	(174)
Xpert MTB RIF, smear-positive TB cases	0.980 (0.010)		beta	(42)
Xpert MTB RIF, smear-negative TB cases	0.680 (0.090)		beta	(42)
Smear microscopy (2 slides), HIV-positive	0.446 (0.036)		beta	(1)
Smear microscopy (2 slides), HIV-negative	0.723 (0.015)		beta	(1)
Mycobacterial culture	1 (-)			model assumption
Clinical diagnosis	0.160 (0.073)	0.44 (0.096)	beta	(1)
specificity for diagnosing pulmonary TB (SE)				
TB-LAMP	0.984 (0.005)		beta	(174)
Xpert MTB RIF	0.980 (0.010)		beta	(42)
Smear microscopy (2 slides)	1 (-)			model assumption
Mycobacterial culture	1 (-)			model assumption
Clinical diagnosis	0.642 (0.009)	0.869 (0.002)	beta	(1)
sensitivity for detecting rifampicin-resistance (SE)				
Xpert MTB RIF	0.950 (0.050)		beta	(42)

Conventional drug susceptibility testing (MGIT DST)	1 (-)			model assumption
specificity for detecting rifampicin-resistance (SE)				
Xpert MTB RIF	0.980 (0.010)		beta	(42)
Drug susceptibility testing	1 (-)			model assumption
Cost parameters (TB treatment) ***US$2014 (min, max)***				
AFB Smear ZN (2 smears)	3.54 (3.01, 4.07)	2.33 (1.98, 2.68)	triangular	(176)
TB-LAMP (national weighted average)	16.01 (13.61, 18.41)	14.42 (12.26, 16.59)	triangular	(176)
Xpert (at intermediate level only)	17.37 (14.76, 19.97)	14.63 (12.44, 17.84)	triangular	(176)
Culture (MGIT)	17.11 (14.54, 19.68)	12.67 (10.77, 14.57)	triangular	(176)
Culture DST (MGIT)	18.79 (15.97, 21.61)	21.41 (18.20, 24.62)	triangular	(176)
Clinical Diagnosis	4.49 (3.82, 5.17)	8.47 (7.20, 9.75)	triangular	WHO-CHOICE (177)
First line CAT 1 treatment	243 (207-279)	354 (301-407)	triangular	(178)
First line CAT 2 treatment	323 (275-372)	471 (400-541)	triangular	(178)
MDR-TB treatment with standardized second line regimen	1281 (1035, 1400)	3407 (2896, 3918)	triangular	(178)
DALY parameters - ***DALYs averted (min, max)***				
HIV + SS-	18.76 (17.67, 18.88)	18.76 (17.67, 18.88)	triangular	For underlying assumptions see S4
HIV - SS-	18.37 (16.03, 19.51)	20.66 (17.19, 22.13)	triangular	For underlying assumptions see S4
HIV + SS+	18.82 (17.71, 18.90)	18.82 (17.71, 18.90)	triangular	For underlying assumptions see S4
HIV - SS+	20.14 (17.59, 22.54)	23.32 (20.44, 25.18)	triangular	For underlying assumptions see S4

Table 1. Model inputs: cohort composition and diagnostic parameters, by country

For patient treatment outcomes, we assumed three possible states – cure, failure, and death – where these treatment outcome probabilities were differentiated by initial patient category (new vs. previously treated) and treatment regimen (first-line category I or II, second-line) based on the data available from the published literature. With regard to drug resistance, the model assumed that there were only two states: drug susceptible and MDR-TB. All cultures and DST are assumed to be on liquid culture media using Mycobacterium growth indicator tube (MGIT) and Bactec 960 equipment. Patients awaiting conventional DST results are started on first-line treatment, and switched to second-line treatment after 12 weeks if the DST shows MDR-TB.

DALYs were calculated using the standard DALY formula with 3% discount rate (69) and were based on age at presentation as observed in the TB-LAMP demonstration study sites (Vietnam and Malawi). Disability weights were taken from the *Global Burden of Disease Study 2010* (179); survival estimates after TB treatment with and without HIV infection were taken from the literature (see Appendix E for full list and references). The DALY calculations assumed that if TB is left untreated, disability is due to TB, and assumed that all TB patients co-infected with HIV receive lifelong highly active antiretroviral therapy (HAART).

Cost parameters

All diagnostic costs parameters are presented in Table 1 and were based on the results from a separate cost-analysis study conducted at the same study sites (176). Costs of diagnostics include following costs: cost of specimen transport, laboratory process, repeat testing due to contamination, and reporting. The primary per-test cost estimate for TB-LAMP assay was based on a weighted average per-test cost, assuming TB-LAMP is implemented at all peripheral (microscopy) laboratories in each respective country. The weights were based on the assumed distribution of peripheral laboratories in the TB laboratory network with low (60%), medium (30%), and high (10%) workloads. These workload levels were represented by average annual SSM testing volume of 1,000, 3,000, and 5,000 smears. High and low estimates of the weighted average per-test cost were calculated by varying the distribution of laboratories with low (between 50-70%), medium (25-35%), and high (5-15%) workloads.

Costs for treatment of DS and MDR-TB were based on the health services costs for TB treatment reported in a systematic review (178). Malawi and Vietnam were considered as low-income and lower-middle income countries, respectively. These estimates included costs of hospitalization, outpatient visits, anti-TB drugs, and other provider costs (e.g. start-up costs, treatment supervision, staff and training, contact tracing, supplies and transportation). For the first-line category 2 treatment (for those previous TB patients whose diagnostics results indicate DS-TB), we assumed that the cost is 130% (based on proportional increase in treatment duration compared to category 1 treatment) of the DS-TB treatment cost reported in the same review. As we used the perspective of a TB services provider, costs of HAART was not evaluated in our model. All costs were reported in 2014 US$ with local costs converted based on the average UN operational exchange rate for 2014 (180). Where applicable, all capital costs were annualized using a standard discount rate of 3% (69).

Sensitivity analysis

Prior to conducting sensitivity analysis, we first constructed a tornado diagram to determine the order of influence on our primary cost-effectiveness estimate. This was done by varying the parameters with the degree of uncertainty obtained from the literature or other sources. When such data were not available, we varied these values at 75% and 125% of the original parameter estimate used for the primary cost-effectiveness estimate. Next, we performed a series of one-way sensitivity analyses on highly influential variables to assess the robustness of our model results. For diagnostic accuracy parameters (e.g. clinical diagnosis), we varied the sensitivity and specificity in opposite directions across the reported range of accuracy values, given the trade-off between the two values (1). We conducted a probabilistic sensitivity analysis (PSA, Monte Carlo simulation, 10,000 iterations) to generate a 95% confidence interval around our primary cost-effectiveness estimate, based on the uncertainty range of all parameters included in our analysis. This was done by randomly sampling each of the parameters used in our model, based on their respective distributions. (1)

We also conducted scenario analysis to test cost-effectiveness of TB-LAMP "add-on" and "replacement" scenarios to SSM in settings where Xpert is implemented as part of the routine diagnostic algorithm. In the base-case scenario in this analysis, Xpert is an initial test (replacement of SSM) for those at risk of MDR-TB and an add-on test only to new presumptive TB patients who are HIV positive and with negative smear result. Therefore, TB-LAMP scenarios are applicable only in patients who would receive SSM as an initial test (new presumptive TB patients). As Xpert per-test cost varies significantly with GeneXpert equipment utilization, we assumed that Xpert is placed at high throughput laboratories (intermediate level laboratory with daily workload of between 10 and 14 test).

Results

In both Malawi and Vietnam, both TB-LAMP scenarios improved case detection rates to between 74-76% and 88-90%, respectively, compared to the base-case scenario rates of 59% and 82% in each country (Table 2). While TB-LAMP is not capable of detecting DR-TB *per se*, improved diagnostic sensitivity over SSM marginally improved MDR-TB case detection rate as well. This was due to the increased sensitivity of TB-LAMP through which it detected more DR-TB cases. The diagnostic cost per TB case detected ranged between $133-161 for

the TB-LAMP add-on and $131-174 for the replacement strategies, whereas the cost for the base case diagnosis ranged between $70-87. The incremental cost per additional TB case detected between the add-on and replacement strategies against the base case were considerably higher in Vietnam ($1015 vs $1484) than Malawi ($337 vs. $413), mainly attributable to the differences in assumptions on the performance of clinical diagnosis.

Country	Scenario	Total TB case detected	% of TB cases detected	Total MDR case detected	% of MDR cases detected	Total Diagnostic Costs (US$2010)	Diagnostic cost per TB case detected, excluding MDR (US$ 2014)	Incremental Cost per additional TB case detected (US$ 2014)
Malawi	Base case	1155	59%	5	32%	80129	70	-
	LAMP add-on	1496	76%	6	41%	194552	131	337
	LAMP replacement	1455	74%	6	40%	203703	141	413
Vietnam	Base case	1207	82%	34	37%	104748	87	-
	LAMP add-on	1317	90%	37	40%	212508	161	1015
	LAMP replacement	1290	88%	36	39%	223843	174	1484

Table 2. Cohort, cases detected, total cohort costs, and costs per TB case detected

Cost-utility (DALYs) analysis results of the two TB-LAMP strategies are presented in Table 3. The cost per DALY averted increased in both TB-LAMP scenarios versus the base case (from $23 to $27 in Malawi and $35 to $40 in Vietnam). The replacement scenario resulted in slightly lower total DALYs averted, but with higher total cost compared to the add-on scenario. The incremental cost-effectiveness ratio (ICER) for the TB-LAMP replacement for SSM scenario was between $41 and $131, which was higher than that of the add-on scenario at $39 and $123 in Malawi and Vietnam, respectively. Both the add-on and replacement scenarios are well below the willingness to pay (WTP) threshold of 1 times the per capita gross domestic product ($255 – Malawi; $2052 – Vietnam) or gross national income ($250 – Malawi; $1890 – Vietnam) (181).

Country	Scenario	Total Cost	Total DALYS	Cost per DALY	ICER compared to base case, mean	Monte Carlo Simulation ICER, median (2.5,97.5)
Malawi	Base Case	510606	21800	23		
	LAMP add-on	752275	28073	27	39	41 (29, 61)
	LAMP replacement	753090	27708	27	41	45 (31, 69)
Vietnam	Base Case	935482	27110	35		
	LAMP add-on	1135705	28744	40	123	145 (69, 294)
	LAMP replacement	1138067	28653	40	131	153 (70, 489)

Table 3. Cost per DALY (US$ 2014)

The results of the PSA are also shown in the same table (Table 3) and both strategies remained cost-effective against respective WTP threshold for each country. Figure 2 shows an illustration of the probability of cost-effectiveness of both strategies against a range of the WTP threshold; the curve shows that the add-on strategy is slightly more favorable over the replacement strategy at lower thresholds. Deterministic one and two way sensitivity analyses results also showed that both TB-LAMP strategies were consistently cost-effective with all uncertain ranges tested producing ICER within the WTP threshold, for all settings. Of the deterministic sensitivity analysis results, accuracy of clinical diagnosis most significantly influenced the ICER for both TB-LAMP strategies (Figure 3). Though they remained within the WTP threshold, at lower specificity (minimum 0.64) and higher sensitivity ranges (maximum 0.80), the cost-effectiveness of both TB-LAMP strategies in comparison to the base case reduces, and the ICERs increase by more than 5 times the primary ICER estimate. Changes in per-test cost based on different distributions of low, medium, and high workload peripheral laboratories did not influence the ICER no more than 15% of the primary estimate. A selection of key one or two way sensitivity analyses results are show in Figure 3.

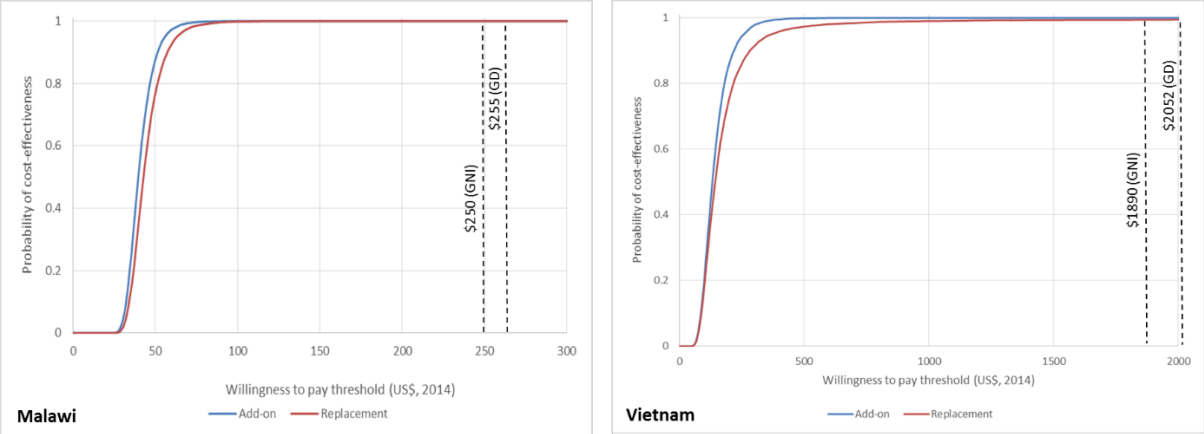


Figure 2. Cost-effectiveness acceptability curves. ICER of "add-on" and "replacement" compared to base case

ICER: Incremental Cost-effectiveness Ratio; GNI: Gross National Income; GDP: Gross Domestic Product

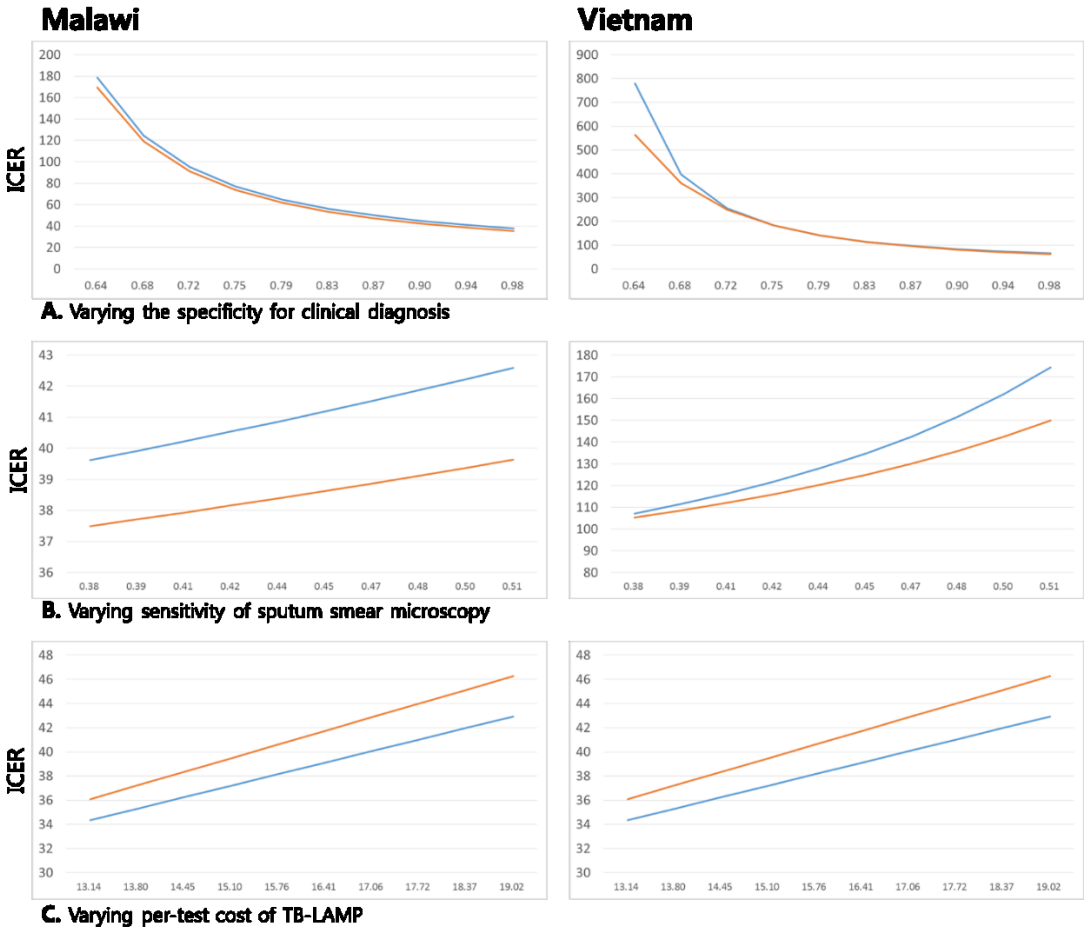


Figure 3. Selected sensitivity analyses results (blue line: replacement strategy; orange line: add-on strategy). Sensitivity of our model for varied A. accuracy of clinical diagnosis, B. sputum smear microscopy in HIV positive patients, and C. per-test cost of TB-LAMP. The sensitivity of clinical diagnosis varies inversely with specificity (range, 8 – 80%). ICERs are reported in US$ 2014.

In our scenario sensitivity analysis, ICER for both the add-on and replacement strategies was higher in settings where Xpert is implemented as a routine diagnostic test compared to the those observed in the non-Xpert base case settings. However, these ICER estimates were within the WTP threshold, suggesting that a combined TB-LAMP and Xpert testing strategy could be a cost-effective strategy in these two countries. When comparing directly between the add-on and replacement strategies, the ICER for the replacement against the add-on strategy was considerably higher than the WTP threshold for Vietnam. This suggest that there is little incremental value for the replacement strategy over the add-on in Vietnam. With respect to case-detection, we found that a combined Xpert and TB-LAMP strategy increased case-detection; in the add-on strategy, the % of cases detected was between 96-97% in both Malawi and Vietnam. These results are presented in Table 4.

Country	Scenario	% of TB cases detected	% of MDR cases detected	Xpert at Intermediate Lab					Xpert at Periperal Lab				
				Diagnostic cost per TB case detected, excluding MDR (US$ 2014)	ICER (Case Detection), US$ 2014	Cost per DALY	ICER (DALY), US$ 2014	ICER (DALY) compared to in addition to, mean	Diagnostic cost per TB case detected, excluding MDR (US$ 2014)	ICER (Case Detection), US$ 2014	Cost per DALY	ICER (DALY), US$ 2014	ICER (DALY) compared to in addition to, mean
Malawi	Base Case (Xpert)	83%	60%	114	-	27	-	-	167	-	30	-	-
	In addition to smear	96%	60%	117	135	28	65	-	163	135	31	65	-
	Replacement of smear	89%	55%	168	924	30	83	97	214	878	33	80	92
Vietnam	Base Case (Xpert)	93%	43%	94	-	43	-	-	106	-	38	-	-
	In addition to smear	97%	43%	150	1485	46	126	-	163	1487	42	159	-
	Replacement of smear	97%	42%	163	1682	46	144	7078	176	1680	42	159	7105

Table 4. Scenario sensitivity analysis. TB-LAMP add-on and replacement strategy compared against base case with Xpert as a routine diagnostic test (HIV positive and smear negative, previously treated TB patients at risk of MDR-TB). Per-test cost for Xpert varied according to the placement level in the health care system (Xpert at Intermediate Lab vs. at Peripheral Lab) where costs were higher for Xpert tested at peripheral Lab.

Discussion

In this decision analytic model, we examined the cost effectiveness of two alternative TB-LAMP strategies to SSM for settings where Xpert is not used as a routine diagnostic test in Malawi and Vietnam. Our model suggests that the use of the TB-LAMP assay increased the TB case detection rate (by 6-15%) and averted DALYs (by 6-28%). Comparing our ICER estimates to the WHO WTP threshold levels (GDP per capita), both strategies were highly cost-effective compared to the base case of using SSM followed by clinical diagnosis (69). We found that there were minimal differences in cost-effectiveness between the two TB-LAMP strategies, though the cost-effectiveness of the add-on strategy was more favorable in both settings.

Our sensitivity analyses, including the PSA, demonstrated robustness of our model results where all analyses results showed that the use of TB-LAMP assay in Malawi and Vietnam is cost-effective. As with other cost-effectiveness analyses of molecular diagnostic tests (e.g. Xpert) (1, 72, 182, 183), our ICER estimate was highly dependent on the accuracy of clinical diagnosis; however, our estimates remained within the WTP threshold levels in both countries. Other factors, such as per-test cost of TB-LAMP assay, increased performance of SSM (reflective of settings using fluorescent microscopy methods), or costs of TB and MDR-TB treatment, that varied in our analyses did not influence our cost-effectiveness assessment of TB-LAMP assay. In regards to the costs, diagnostic cost estimates were based on bottom-up micro costing method using time and motion study (184). Per-test cost estimates reflected a complete diagnostic process and were calculated for varied levels of operational characteristics, where these estimates were most influenced by the laboratory's diagnostic workload (cost analysis study).

In our scenario sensitivity analysis, we found that TB-LAMP assay can also potentially be cost-effective in settings where Xpert tests are being offered under the routine diagnostic algorithm. Compared to the non-Xpert scenario (primary analysis), ICER for the two TB-LAMP strategies were considerably higher, with ICER estimates > 70% of our primary estimate. In the case of Vietnam, TB-LAMP replacement strategy was not found to be cost-effective when compared directly to the add-on strategy as the baseline, with ICER beyond the WTP threshold (ICER 7078). These findings suggest that TB-LAMP may be most cost-

effective when used in addition to SSM, particularly for presumptive TB patients whose SSM results are negative.

Our findings are subject to several limitations. First, our model is a greatly simplified illustration of TB diagnosis and management, which in reality, is a much complex and dynamic process (72). Our assessment of the varied accuracy of clinical diagnosis is reflective of only a small part of the realities faced at each clinic. For example, as found in a randomized control study of Xpert in South Africa (66), any potential benefit of improved diagnostic performance can be negated by wide practice of empirical treatment. Furthermore, as shown in our results, diagnostics are a small (less than 30% of the total cost) component of the complete health care process of patients investigated and treated for TB. Likewise, relying on diagnostic accuracy parameters to project costs and patient relevant outcomes (e.g. DALYs) is likely to over-estimate potential cost-effectiveness of TB-LAMP.

Secondly, due to the lack of relevant data in our study settings, we did not assess the impact of reduction in diagnostic delay on treatment, patient health outcomes, and TB transmission. Additionally, our analysis was limited to the health systems perspective and TB diagnosis and treatment at the public sector. A modeling study for India incorporating these factors has shown that a public-private mixed implementation strategy of a new diagnostic test for TB (e.g. Xpert) can considerably influence the impact of new diagnostics (141). Furthermore, rapid NAATs have been shown to considerably reduce both diagnostic and therapeutic delays (Sohn et al., unpublished systematic review results) and may potentially influence healthcare seeking behaviors. Moreover, given the evidence that delays experienced by presumptive TB patients were attributable to potentially unnecessary repeated visits at the primary care level (6, 32, 85), it is likely that inclusion of patient costs and assessment in the complete societal perspective would further favor rapid diagnostic tests such as TB-LAMP.

Thirdly, we rely on the annual per capital GDP (or GNI) as the WTP threshold per standard practice in evaluating cost-effectiveness when costs are projected against health outcome estimates such as DALYs (69). When these WTP thresholds are compared to ICERs calculated based on more direct and meaningful outcomes of diagnostic test performance (e.g. improved case detection), the same cost-effectiveness conclusion may not hold. Furthermore, it is important to note that a cost-effective strategy does not translate to a strategy that is affordable or sustainable (1, 72). We have addressed these issues in a separate

cost and affordability assessment of TB-LAMP and found that nationwide implementation of TB-LAMP at all peripheral level laboratories can account for a considerable portion (more than 10%) of the total annual TB budget (176). In the case of Malawi, these costs were approximately 17% of the total projected budget for the national TB control program. Considering that a large proportion of the annual national TB budgets in high TB-burden, resource-limited settings rely on external donor funding and highly setting specific, it is important to evaluate potential implications of the implementation of TB-LAMP on resource mobilization and financial sustainability post-implementation prior to making national level policy decisions.

Conclusion

Our model results demonstrate that TB-LAMP is potentially a cost-effective alternative to the base case of SSM + clinical diagnosis in settings where Xpert is not available or is highly costly to operate. Of the two strategies evaluated, using LAMP as an add-on test to patients with negative SSM results was consistently more cost-effective compared to the replacement strategy in all levels of our analyses. However, given TB-LAMP's lack of capacity to detect DR-TB, financial resource constraints at each country, and emerging POC molecular tests for TB, policy makers must cautiously evaluate operational and financial feasibility prior to introducing this technology to their respective countries.

Chapter 6: Summary & Conclusion

6.1 Summary of results

Diagnostic research is an important area of study that requires a highly multi-disciplinary approach to evaluate a test's accuracy, applicability and feasibility, and clinical and economic impact. Although Xpert MTB/RIF (Xpert) assay has revolutionized the field of TB diagnostics research, with some studies now evaluating impact beyond test accuracy, the vast majority of evidence is still limited to the realm of diagnostic performance. As TB point-of-care (POC) molecular tests emerge from development to field testing, it is imperative to address current knowledge gaps in the diagnostics field, and ultimately provide guidance on future research methods that could better inform policy decisions. Therefore, my thesis employed a multi-disciplinary approach in evaluating current molecular diagnostic tests for TB in three key areas of diagnostic research – synthesis of the evidence regarding the impact of NAATs on reducing time delays in diagnosis and treatment of TB, assessment of test performance and feasibility of POC Xpert test in a low incidence setting, and economic evaluation of a novel molecular assay implemented as a replacement or add-on test to sputum smear microscopy at the peripheral microscopy centers.

In my first manuscript (Chapter 3), I systematically reviewed the comparative impact of the WHO-approved nucleic acid amplification tests (NAAT) on improving diagnostic and therapeutic delays for TB. First, I developed a revised conceptual framework in defining time delays in TB patient care. This was used to qualitatively and quantitatively summarize diagnostic and therapeutic delays of diagnosis and management of TB patients in situations with and without use of NAATs. This review of 39 identified studies found that time delays in diagnosis and treatment of both drug-susceptible (DS) and drug-resistant (DR) forms of TB were reduced when NAATs were used for diagnosis and clinical management of TB patients. However, the magnitude and significance of the delay assessment varied considerably in sub-group analyses due to heterogeneity observed amongst the studies included. While studies represented a wide range of settings, a large proportion of this research was concentrated in the African region, mostly in South Africa, limiting generalizability of the study findings. Considerable variations in methodological approaches exist in publications evaluating time delays, from discrepancies in defining time components to statistical methods used in evaluation of time delay impact; this was likely because the majority of studies' primary

objectives were to test performance evaluation, and time delay assessment was evaluated as a secondary outcome; additionally, this assessment was mostly based on retrospective reviews of medical chart and records. Lastly, a quality assessment framework was proposed to not only evaluate the quality of time estimates reported, but also to provide guidance on future research evaluating time delays in clinical care. Based on the proposed criteria, I found that a majority of studies lacked well-defined frameworks and methods for evaluating time delays, limiting comparability and generalizability of study findings.

In my second manuscript (Chapter 4), I assessed diagnostic performance and feasibility of a POC Xpert testing strategy in a cohort of patients suspected of TB infection in a specialized TB clinic in Montreal, Canada, that is, in a high resource, low TB incidence setting. In contrast to earlier findings, I found that Xpert's performance and potential impact in such settings was relatively limited, largely as a result of its lower sensitivity in a context of less prevalent disease and highly efficient diagnostic algorithms and clinical operations. More specifically, the considerably low sensitivity of Xpert in diagnosing smear-negative TB undermines its applicability as a rapid rule-out POC test and impact on patient care, given the availability of culture results and highly trained physicians in the diagnosis and care of TB patients. Altogether, this study shows that in low TB incidence settings, Xpert results should be made use of with caution for clinical management of TB patients, and all Xpert results be confirmed by conventional routine diagnostics.

In my third manuscript, I used decision analytic modeling methods to evaluate the cost-effectiveness of TB-LAMP assay, a rapid and simple molecular test that can be implemented in peripheral microscopy centers in resource limited settings. Two distinct diagnostic strategies – TB-LAMP assay used as a replacement for or as an add-on test to sputum smear microscopy using Zhiel Neelsen (ZN) staining (SSM) – were evaluated against a baseline strategy of SSM followed by clinical diagnosis in patients with negative SSM in two low MDR-TB prevalence settings, Malawi and Vietnam. The model results suggest that use of TB-LAMP assay to supplement or replace SSM in these settings improves case-detection and health outcomes, measured by the number of DALYs averted. Evaluating the incremental cost-effectiveness ratio of both intervention strategies against the WHO willingness to pay (WTP) threshold implies that both TB-LAMP strategies are highly cost-effective, with our range of sensitivity analysis results indicating robustness of this initial conclusion. As modeling studies are a way of identifying important gaps in research, this study demonstrates

the importance of understanding the impact of clinical diagnosis, health system operations and effectiveness, and patients' healthcare seeking behaviors on cost and cost-effectiveness of novel diagnostics and interventions in resource limited settings.

6.2 Strengths and Limitations

My systematic review (manuscript 1) is the first review to assess, post-WHO endorsement, the impact of NAATs on reducing time delays in diagnosis and treatment of TB. Key strengths of this study include use of an extensive search strategy and a critical and complete assessment of delays (both qualitative and quantitative) associated with both DS and DR-TB, which is based on a detailed conceptual framework evaluating time delay impact of TB diagnostic tests. The conceptual framework for defining time delay components and assessing quality of time delay estimates used in this study can serve as a guideline for future research evaluating time delay impact. Conversely, lack of individual patient data on time delay estimates limit methodologies used for quantitative assessment. Thus, while the quality assessment criteria is based on sound epidemiological and statistical principles, it was not developed based on validated tools.

My primary research study (manuscript 2) is the first study in Canada of Xpert MTB/RIF assay and one of the first studies in a low TB burden setting. This study was not limited to evaluating diagnostic performance, but instead also assessed potential impact beyond diagnosis, particularly in reducing diagnostic and treatment delays. It is also one of the first studies to explore applicability and feasibility of POC testing strategies using Xpert. Key limitations of this study include the low generalizability of findings to settings that differ from a high resource, highly specialized TB clinic. Furthermore, due to the lack of Xpert's regulatory approval for use in Canada at the time of this study, the impact assessment was based on a hypothetical 'would have been' scenario; additionally, other impact further downstream of diagnosis was not evaluated.

In my final manuscript (manuscript 3), I developed a decision analytic model that was based on a well-established diagnostic cost-effectiveness analysis model (I actively contributed in the development of the original model). It is the first economic evaluation of a TB-LAMP assay, which was supplemented with setting-specific epidemiologic data that represented two distinct types of high TB-burden, low resource settings. An additional strength of this study is its use of highly detailed economic cost analysis results that reflect a range of different test roll-out

scenarios. Furthermore, the robustness of the main study results are supported by a wide range of sensitivity analyses that include probabilistic and scenario sensitivity analyses. There are several limitations in this study. First, my decision analytic model is a much simplified illustration of the complex TB diagnosis management that is observed in high TB-burden, resource poor settings. Second, it is also a model that predicts clinical and patient impact largely based on diagnostic performance (sensitivity and specificity). Lastly, the model did not incorporate benefits and impact of reduced timed delays on treatment and patient health outcomes, or disease transmission due to lack of relevant setting specific data.

6.3 Implications & directions for future research

The range of studies that comprise of my doctoral thesis provide critical and comprehensive evidence of the impact of new diagnostic strategies beyond test accuracy; this kind of evidence is now required for policy and guideline development for bodies such the World Health Organization (WHO). My studies also help improve research methodologies used to evaluate tests' impact beyond accuracy, and highlight important areas and directions of future research in TB diagnostics.

Specifically, my systematic review (manuscript 1) is the first review that critically assessed the impact of post WHO-policy implementation of new diagnostic tools on effective reduction of time delays in TB. My review highlights important limitations in current research methods in evaluating the effect of diagnostic interventions on time delays and provides an improved conceptual framework so that future research in this area can produce comparable and high quality evidence. Subsequently, the qualitative and quantitative evidence generated from this study support the important idea that limited improvement in time delays in diagnosis and treatment of TB is highly attributable to factors associated with health systems operations as well as physician and patient behaviors. Likewise, it is clear that future TB diagnostic research must evaluate these factors so that the impact of novel molecular TB diagnostics may be optimized.

Secondly, my primary research study addressed an important gap in research surrounding test accuracy and impact of Xpert MTB/RIF used as a POC test in low incidence settings. Its documented limited impact and performance underscore the recommendation in the Canadian TB standards that use of Xpert should not replace current laboratory practice for diagnosis of TB in settings with low incidence and highly effective clinical operations. Assessment of

feasibility of POC testing highlighted important considerations for biosafety and logistical issues that future POC tests must consider, and provides directions for future research into the applicability of Xpert in more remote Canadian settings where testing capacity is highly limited. I did not evaluate cost and cost-effectiveness of Xpert or similar technology on TB patient care in these study settings; however, these are critical areas of research to follow up on that will provide evidence for policy development in high resource settings.

In my final research manuscript (manuscript 3), I highlight the importance of health economic impact evaluations as part of diagnostic research where evaluation of the epidemiological impact of an intervention against economic costs is an important prerequisite for sound policy decision making. The results from this research provided important health economic evidence to the WHO policy and guideline development for TB-LAMP assay. Cost-effectiveness analysis alone should not be considered as sufficient evidence for policy decision-making process; although not part of the thesis, my cost-effectiveness analysis was based on a rigorous assessment of the economic cost of TB-LAMP implementation strategies, and additional work on affordability of intended implementation strategies was also provided to the WHO as part of a comprehensive economic evaluation package. This study demonstrated that further evaluation of the effectiveness of clinical diagnosis and methods for incorporating health systems operational factors (public, private mix) is needed to maximize the benefits of rapid molecular diagnostics in low resource settings.

6.4 Conclusion

This doctoral thesis, comprised of three studies, utilized multi-disciplinary research methods and approaches to highlight the variety of diagnostic research processes currently in practice. They range from evaluation of test accuracy to various impact assessments, including assessment of cost and cost-effectiveness, with an emphasis on complete knowledge translation into practice and policy decision making. Results and findings from my research indicate that more attention is needed in understanding operational, clinical, and patient factors that affect delays in diagnosis and treatment of TB, so that future evaluation and implementation diagnostic of tools for TB can be optimized to reduce burden for resource allocation. Furthermore, it demonstrates that future TB diagnostics research should not only thoroughly evaluate test accuracy, but also be rigorous in evaluating potential impact beyond diagnostic accuracy. Assessment of applicability and feasibility of process innovation is an

important area of research that can synergistically effect improvement in TB care and the fight against ending TB. Finally, evaluation of cost effectiveness should be supplemented by sound epidemiological and costing data to accurately reflect specific settings and intended health interventions. These research methods should not be limited to generating evidence, but should be employed to identify gaps in research that can better support evidence-based medical decision making processes.

Back Matter

References

1. Vassall A, van Kampen S, Sohn H, Michael JS, John KR, den Boon S, et al. Rapid diagnosis of tuberculosis with the Xpert MTB/RIF assay in high burden countries: a cost-effectiveness analysis. PLoS Med. 2011;8(11):e1001120.
2. World Health Organization. Global Tuberculosis Report 2015: World Health Organization; 2015.
3. Public health Agency of Canada. Tuberculosis in Canada 2013 - Pre-release. . Ottawa: Public Health Agency of Canada, 2015.
4. World Health Organization. New technologies for tuberculosis control: a framework for their adoption, introduction and implementation [WHO/HTM/STB/2007.40]. Geneva: World Health Organization, 2007.
5. Storla DG, Yimer S, Bjune GA. A systematic review of delay in the diagnosis and treatment of tuberculosis. BMC Public Health. 2008;8:15.
6. Sreeramareddy CT, Kishore PV, Menten J, Van den Ende J. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. BMC infectious diseases. 2009;9(1):91.
7. Keeler E, Perkins MD, Small P, Hanson C, Reed S, Cunningham J, et al. Reducing the global burden of tuberculosis: the contribution of improved diagnostics. Nature. 2006;444 Suppl 1:49-57.
8. Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM, Jr., Dye C, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. Proc Natl Acad Sci U S A. 2009;106(33):13980-5.
9. Perkins MD, Roscigno G, Zumla A. Progress towards improved tuberculosis diagnostics for developing countries. Lancet. 2006;367(9514):942-3.
10. Pai M, Schito M. Tuberculosis diagnostics in 2015: landscape, priorities, needs, and prospects. Journal of Infectious Diseases. 2015;211(suppl 2):S21-S8.
11. World Health Organization. Policy statement. Molecular line probe assays for rapid screening of patients at risk of multidrug-resistant tuberculosis (MDR-TB). URL: http://www.who.int/tb/features_archive/policy_statement.pdf Geneva: World Health Organization; 2008 [cited 2008]. Available from: http://www.who.int/tb/dots/laboratory/policy/en/index4.html.
12. World Health Organization. Policy statement: automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. 2011.
13. Raviglione M, Marais B, Floyd K, Lönnroth K, Getahun H, Migliori GB, et al. Scaling up interventions to achieve global tuberculosis control: progress and new developments. The Lancet. 2012;379(9829):1902-13.
14. Pai M, Minion J, Steingart K, Ramsay A. New and improved tuberculosis diagnostics: evidence, policy, practice, and impact. Current opinion in pulmonary medicine. 2010;16(3):271-84.
15. Small PM, Pai M. Tuberculosis diagnosis—time for a game change. New England Journal of Medicine. 2010;363(11):1070-1.
16. World Health Organization. Toman's tuberculosis: case detection, treatment, and monitoring (second edition). Geneva: World Health Organization, 2004.

17. World Health Organization. Global tuberculosis control 2010. Geneva: World Health Organization, 2010.
18. Dye C, Williams BG. The population dynamics and control of tuberculosis. Science. 2010;328(5980):856-61.
19. Public Health Agency of Canada Tuberculosis in Canada 2009 - Pre-release. Ottawa: Public Health Agency of Canada, 2010.
20. Ellis E, Sauve L, Phypers M, Sheardown C, Allegakone M. Tuberculosis in Canada. Ontario: Public Health Agency of Canada, 2002.
21. Macdonald N, Hebert PC, Stanbrook MB. Tuberculosis in Nunavut: a century of failure. CMAJ. 2011.
22. Steingart KR, Ramsay A, Pai M. Optimizing sputum smear microscopy for the diagnosis of pulmonary tuberculosis. Expert Rev Anti Infect Ther. 2007;5(3):327-31.
23. 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.
24. Perkins MD, Small PM. Admitting defeat. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2006;10(1):1.
25. Public Health Agency of Canada & Canadian Lung Association. Canadian Tuberculosis Standards, 6th Edition. Ottawa: Public Health Agency of Canada & Canadian Lung Association, 2007.
26. Squire SB, Belaye AK, Kashoti A, Salaniponi FM, Mundy CJ, Theobald S, et al. 'Lost' smear-positive pulmonary tuberculosis cases: where are they and why did we lose them? The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2005;9(1):25-31.
27. Davies PD, Pai M. The diagnosis and misdiagnosis of tuberculosis. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2008;12(11):1226-34.
28. Perkins MD. New diagnostic tools for tuberculosis. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2000;4(12 Suppl 2):S182-8.
29. Greco S, Girardi E, Navarra S, Saltini C. The current evidence on diagnostic accuracy of commercial based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis. Thorax. 2006;61(9):783-90.
30. Palomino JC. Molecular detection, identification and drug resistance detection in Mycobacterium tuberculosis. FEMS Immunol Med Microbiol. 2009;56(2):103-11.
31. Sreeramareddy CT, Panduru KV, Menten J, Van den Ende J. Time delays in diagnosis of pulmonary tuberculosis: a systematic review of literature. BMC infectious diseases. 2009;9:91.
32. Storla DG, Yimer S, Bjune GA. A systematic review of delay in the diagnosis and treatment of tuberculosis. BMC public health. 2008;8(1):15.
33. Yimer S, Bjune G, Alene G. Diagnostic and treatment delay among pulmonary tuberculosis patients in Ethiopia: a cross sectional study. BMC infectious diseases. 2005;5(1):112.
34. Sreeramareddy CT, Qin ZZ, Satyanarayana S, Subbaraman R, Pai M. Delays in diagnosis and treatment of pulmonary tuberculosis in India: a systematic review. The international journal of tuberculosis and lung disease: the official journal of the International Union against Tuberculosis and Lung Disease. 2014;18(3):255-66.
35. Menzies D. Screening immigrants to Canada for tuberculosis: chest radiography or tuberculin skin testing? CMAJ. 2003;169(10):1035-6.

36. 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;353(9151):444-9.
37. Hernandez-Garduno E, Cook V, Kunimoto D, Elwood R, Black W, FitzGerald J. Transmission of tuberculosis from smear negative patients: a molecular epidemiology study. Thorax. 2004;59(4):286-90.
38. Millen SJ, Uys PW, Hargrove J, Van Helden PD, Williams BG. The effect of diagnostic delays on the drop-out rate and the total delay to diagnosis of tuberculosis. PLoS One. 2008;3(4):e1933.
39. MacPherson P, Houben RM, Glynn JR, Corbett EL, Kranzer K. Pre-treatment loss to follow-up in tuberculosis patients in low-and lower-middle-income countries and high-burden countries: a systematic review and meta-analysis. Bulletin of the World Health Organization. 2014;92(2):126-38.
40. Dowdy DW, Davis JL, den Boon S, Walter ND, Katamba A, Cattamanchi A. Population-level impact of same-day microscopy and Xpert MTB/RIF for tuberculosis diagnosis in Africa. PloS one. 2013;8(8):e70485.
41. Boehme CC, Nabeta P, Hillemann D, Nicol MP, Shenai S, Krapp F. Rapid molecular detection of tuberculosis and rifampin resistance. N Engl J Med. 2010;363.
42. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev. 2014;1.
43. Banada PP, Sivasubramani SK, Blakemore R, Boehme C, Perkins MD, Fennelly K, et al. Containment of bioaerosol infection risk by the Xpert MTB/RIF assay and its applicability to point-of-care settings. Journal of clinical microbiology. 2010;48(10):3551-7.
44. Helb D, Jones M, Story E, Boehme C, Wallace E, Ho K, et al. Rapid detection of Mycobacterium tuberculosis and rifampin resistance by use of on-demand, near-patient technology. Journal of clinical microbiology. 2010;48(1):229-37.
45. World Health Organization. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. 2013.
46. Wold Health Organization. TBXpert Prject 2013 [January 2015]. Available from: http://who.int/tb/publications/TBXpert_briefing_note.pdf?ua=1.
47. Foundation for Innovative New Diagnostics. Price for Xpert® MTB/RIF and FIND country list 2013 [cited 2013 November]. Available from: http://www.finddiagnostics.org/about/what_we_do/successes/find-negotiated-prices/xpert_mtb_rif.html.
48. World Health Organization. WHo monitoring of Xpert MTB/RIF roll-out 2015 [cited 2015 January]. Available from: http://apps.who.int/tb/laboratory/xpertmap/.
49. Pantoja A, Fitzpatrick C, Vassall A, Weyer K, Floyd K. Xpert MTB/RIF for diagnosis of tuberculosis and drug-resistant tuberculosis: a cost and affordability analysis. Eur Respir J. 2013;42(3):708-20.
50. UNITAID. Tuberculosis Diagnostics Technology and Market Landscape, 4th Edition. World Health Organization, 2015.
51. Goodman C. HTA 101: Introductin to Health Technology Assessment. Bethesda, MD: National Library of Medicine (US), 2014.
52. Fanourgiakis J, Kanoupakis E. Health technology assessment (HTA): a brief introduction of history and the current status in the field of cardiology under the economic crisis. Journal of Evidence-Based Medicine. 2015;8(3):161-4.

53. Lijmer JG, Leeflang M, Bossuyt PM. Proposals for a phased evaluation of medical tests. Medical Decision Making. 2009;29(5):E13-E21.
54. Freedman LS. Evaluating and comparing imaging techniques: a review and classification of study designs. The British journal of radiology. 1987;60(719):1071-81.
55. Zweig MH, Robertson EA. Why we need better test evaluations. Clinical chemistry. 1982;28(6):1272-6.
56. Sackett D, Haynes R. The architecture of diagnostic research. Bmj. 2002;324(7336):539-41.
57. Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. Bmj. 2002;324(7335):477-80.
58. Dukic V, Gatsonis C. Meta-analysis of Diagnostic Test Accuracy Assessment Studies with Varying Number of Thresholds. Biometrics. 2003;59(4):936-46.
59. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Bmj. 2008;336(7653):1106-10.
60. Brunet L, Minion J, Lienhardt C, Pai M. Mapping the landscape of tuberculosis diagnostic research. American journal of respiratory and critical care medicine. 2010;181:A2255.
61. Cobelens F, van den Hof S, Pai M, Squire SB, Ramsay A, Kimerling ME. Which new diagnostics for tuberculosis, and when? Journal of Infectious Diseases. 2012;205(suppl 2):S191-S8.
62. Lord SJ, Irwig L, Bossuyt PM. Using the principles of randomized controlled trial design to guide test evaluation. Med Decis Making. 2009;29(5):E1-E12.
63. Lord SJ, Irwig L, Simes RJ. When is measuring sensitivity and specificity sufficient to evaluate a diagnostic test, and when do we need randomized trials? Ann Intern Med. 2006;144(11):850-5.
64. Pai M, Minion J, Steingart K, Ramsay A. New and improved tuberculosis diagnostics: evidence, policy, practice, and impact. Curr Opin Pulm Med. 2010;16(3):271-84.
65. Biesheuvel CJ, Grobbee DE, Moons KG. Distraction from randomization in diagnostic research. Ann Epidemiol. 2006;16(7):540-4.
66. Theron G, Zijenah L, Chanda D, Clowes P, Rachow A, Lesosky M, et al. Feasibility, accuracy, and clinical effect of point-of-care Xpert MTB/RIF testing for tuberculosis in primary-care settings in Africa: a multicentre, randomised, controlled trial. The Lancet. 2014;383(9915):424-35.
67. Lessells RJ, Cooke GS, Newell M-L, Godfrey-Faussett P. Evaluation of tuberculosis diagnostics: establishing an evidence base around the public health impact. Journal of Infectious Diseases. 2011;204(suppl 4):S1187-S95.
68. Schumacher S. Estimating the impact of tuberculosis diagnostics on patient outcomes: methodological challenges and opportunities. Montreal, Canada: McGill University; 2015.
69. Edejer TT-T. Making choices in health: WHO guide to cost-effectiveness analysis: World Health Organization; 2003.
70. Canadian Agency for Drugs and Technologies in Health. Guidelines for the Economic Evaluation of Health Technologies: Canada. Ottawa, Canada: 2006.
71. Earnshaw J, Lewis G. NICE guide to the methods of technology appraisal. Pharmacoeconomics. 2008;26(9):725-7.
72. Dowdy DW, Cattamanchi A, Steingart KR, Pai M. Is scale-up worth it? Challenges in economic analysis of diagnostic tests for tuberculosis. PLoS Med. 2011;8(7):e1001063.

73. Bill & Melinda Gates Foundation. The Gates Reference Case: What it is, why it's important, and how to use it London: NICE International; 2014 [cited 2015 June 4]. Available from: https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-International/projects/Gates-Reference-case-what-it-is-how-to-use-it.pdf.
74. World Health Organization, Sachs J. Macroeconomics and health: investing in health for economic development: report of the Commission on Macroeconomics and Health: WHO; 2001.
75. Oxlade O, Pinto M, Trajman A, Menzies D. How methodologic differences affect results of economic analyses: a systematic review of interferon gamma release assays for the diagnosis of LTBI. PloS one. 2013;8(3):e56044.
76. Uys PW, Warren RM, Van Helden PD. A threshold value for the time delay to TB diagnosis. PloS one. 2007;2(8):e757.
77. Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC. Consensus statement. Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. WHO Global Surveillance and Monitoring Project. JAMA. 1999;282(7):677-86.
78. Ling DI, Zwerling AA, Pai M. GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: a meta-analysis. Eur Respir J. 2008;32(5):1165-74.
79. Cohen D, Corbett E. Evidence supports TB test, so what now? Cochrane Database Syst Rev. 2013;2:ED000051.
80. Ramsay A, Steingart KR, Pai M. Assessing the impact of new diagnostics on tuberculosis control [Editorial]. The International Journal of Tuberculosis and Lung Disease. 2010;14(12):1506-7.
81. Schunemann HJ, Oxman AD, Brozek J, Glasziou P, Bossuyt P, Chang S, et al. GRADE: assessing the quality of evidence for diagnostic recommendations. Evidence-based medicine. 2008;13(6):162-3.
82. Lienhardt C, Espinal M, Pai M, Maher D, Raviglione MC. What research is needed to stop TB? Introducing the TB Research Movement. The Lancet infectious diseases. 2011;8(11):1412.
83. 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 infectious diseases. 2005;5:62.
84. Denkinger CM, Schumacher SG, Boehme CC, Dendukuri N, Pai M, Steingart KR. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. European Respiratory Journal. 2014:erj00078-2014.
85. Li Y, Ehiri J, Tang S, Li D, Bian Y, Lin H, et al. Factors associated with patient, and diagnostic delays in Chinese TB patients: a systematic review and meta-analysis. BMC medicine. 2013;11(1):156.
86. Higgins J. Green S. Cochrane handbook for systematic reviews of interventions version 5.1. 0. The Cochrane Collaboration. 2011;5(0).
87. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Annals of internal medicine. 2009;151(4):264-9.
88. Li T, Vedula SS, Hadar N, Parkin C, Lau J, Dickersin K. Innovations in Data Collection, Management, and Archiving for Systematic ReviewsData Collection Innovations for Systematic Reviews. Annals of internal medicine. 2015;162(4):287-94.
89. Lippincott CK, Miller MB, Popowitch EB, Hanrahan CF, Van Rie A. Xpert (R) MTB/RIF Shortens Airborne Isolation for Hospitalized Patients with Presumptive Tuberculosis in the United States. Clinical Infectious Diseases. 2014;59(2):21.

90. Chaisson LH, Roemer M, Cantu D, Haller B, Millman AJ, Cattamanchi A, et al. Impact of GeneXpert MTB/RIF assay on triage of respiratory isolation rooms for inpatients with presumed tuberculosis: a hypothetical trial. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2014;59(10):1353-60.
91. Dlamini-Mvelase NR, Werner L, Phili R, Cele LP, Mlisana KP. Effects of introducing Xpert MTB/RIF test on multi-drug resistant tuberculosis diagnosis in KwaZulu-Natal South Africa. BMC infectious diseases. 2014;14:442.
92. Fontela P, Pai N, Dendukuri N, Schiller I, XVIII IEA World Congress of Epidemiology and the VII Brazilian Congress of Epidemiology B, 21-24 September, 2008. , Ramsay A, et al. Quality of diagnostic accuracy studies: evaluation using QUADAS and STARD standards. XVIII IEA World Congress of Epidemiology and the VII Brazilian Congress of Epidemiology, 20-24 September, 2008 Porto Alegre, Brazil2008.
93. Zhang B, Zhang Y. Mann-Whitney U test and Kruskal-Wallis test should be used for comparisons of differences in medians, not means: Comment on the article by van der Helm-van Mil et al. Arthritis & Rheumatism. 2009;60(5):1565-.
94. Wan X, Wang W, Liu J, Tong T. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. BMC medical research methodology. 2014;14(1):135.
95. Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. Effect Sizes Based on Means. Introduction to Meta-Analysis: John Wiley & Sons, Ltd; 2009. p. 21-32.
96. Borenstein M, Hedges L, Rothstein H. Meta-analysis: fixed effect vs. random effects. Meta-Analysis com. 2007.
97. Hardy RJ, Thompson SG. Detecting and describing heterogeneity in meta-analysis. Statistics in medicine. 1998;17(8):841-56.
98. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in medicine. 2002;21(11):1539-58.
99. Borenstein M, Hedges LV, Higgins J, Rothstein HR. Identifying and quantifying heterogeneity. Introduction to Meta-analysis. 2009:107-25.
100. Auld SC, Moore BK, Killam WP, Eng B, Nong K, Pevzner EC, et al. Rollout of Xpert MTB/RIF in northwest Cambodia for the diagnosis of tuberculosis among PLHA. Public Health Action. 2014;4(4):216-21.
101. Mupfumi L, Makamure B, Chirehwa M, Sagonda T, Zinyowera S, Mason P, et al., editors. Impact of Xpert MTB/RIF on antiretroviral therapy-associated tuberculosis and mortality: a pragmatic randomized controlled trial. Open forum infectious diseases; 2014: Oxford University Press.
102. Cox HS, Mbhele S, Mohess N, Whitelaw A, Muller O, Zemanay W, et al. Impact of Xpert MTB/RIF for TB diagnosis in a primary care clinic with high TB and HIV prevalence in South Africa: a pragmatic randomised trial. PLoS Med. 2014;11(11):e1001760.
103. Durovni B, Saraceni V, van den Hof S, Trajman A, Cordeiro-Santos M, Cavalcante S, et al. Impact of replacing smear microscopy with Xpert MTB/RIF for diagnosing tuberculosis in Brazil: a stepped-wedge cluster-randomized trial. PLoS Med. 2014;11(12):e1001766.
104. Churchyard GJ, Stevens WS, Mametja LD, McCarthy KM, Chihota V, Nicol MP, et al. Xpert MTB/RIF versus sputum microscopy as the initial diagnostic test for tuberculosis: a cluster-randomised trial embedded in South African roll-out of Xpert MTB/RIF. The Lancet Global Health. 2015;3(8):e450-e7.
105. 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. The lancet. 2011;377(9776):1495-505.

106. Yoon C, Cattamanchi A, Davis JL, Worodria W, den Boon S, Kalema N, et al. Impact of Xpert MTB/RIF testing on tuberculosis management and outcomes in hospitalized patients in Uganda. PLoS One. 2012;7(11):e48599.
107. Sohn H, Aero AD, Menzies D, Behr M, Schwartzman K, Alvarez GG, et al. Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact. Clinical infectious diseases. 2014:ciu022.
108. Davis JL, Kawamura LM, Chaisson LH, Grinsdale J, Benhammou J, Ho C, et al. Impact of GeneXpert MTB/RIF on patients and tuberculosis programs in a low-burden setting. a hypothetical trial. American journal of respiratory and critical care medicine. 2014;189(12):1551-9.
109. Balcells ME, Garcia P, Chanqueo L, Bahamondes L, Lasso M, Gallardo AM, et al. Rapid molecular detection of pulmonary tuberculosis in HIV-infected patients in Santiago, Chile. International journal of tuberculosis and lung disease. 2012;16(10):1349-53.
110. Van Rie A, Page-Shipp L, Hanrahan CF, Schnippel K, Dansey H, Bassett J, et al. Point-of-care Xpert(R) MTB/RIF for smear-negative tuberculosis suspects at a primary care clinic in South Africa. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2013;17(3):368-72.
111. Hanrahan CF, Selibas K, Deery CB, Dansey H, Clouse K, Bassett J, et al. Time to treatment and patient outcomes among TB suspects screened by a single point-of-care xpert MTB/RIF at a primary care clinic in Johannesburg, South Africa. PLoS One. 2013;8(6):e65421.
112. Kwak N, Choi SM, Lee J, Park YS, Lee CH, Lee SM, et al. Diagnostic Accuracy and Turnaround Time of the Xpert MTB/RIF Assay in Routine Clinical Practice. PLoS One. 2013;8(10):e77456.
113. Buchelli Ramirez HL, Garcia-Clemente MM, Alvarez-Alvarez C, Palacio-Gutierrez JJ, Pando-Sandoval A, Gagatek S, et al. Impact of the XpertReg. MTB/RIF molecular test on the late diagnosis of pulmonary tuberculosis. International Journal of Tuberculosis and Lung Disease. 2014;18(4):435-7.
114. Cohen GM, Drain PK, Noubary F, Cloete C, Bassett IV. Diagnostic delays and clinical decision making with centralized Xpert MTB/RIF testing in Durban, South Africa. Journal of acquired immune deficiency syndromes (1999). 2014;67(3):e88-93.
115. Kim SY, Kim H, Kim SY, Ra EK, Joo SI, Shin S, et al. The Xpert(R) MTB/RIF assay evaluation in South Korea, a country with an intermediate tuberculosis burden. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2012;16(11):1471-6.
116. Kim CH, Woo H, Hyun IG, Kim C, Choi JH, Jang SH, et al. A comparison between the efficiency of the Xpert MTB/RIF assay and nested PCR in identifying Mycobacterium tuberculosis during routine clinical practice. Journal of thoracic disease. 2014;6(6):625-31.
117. Omrani AS, Al-Otaibi MF, Al-Ateah SM, Al-Onazi FM, Baig K, El-Khizzi NA, et al. GeneXpert MTB/RIF Testing in the Management of Patients with Active Tuberculosis; A Real Life Experience from Saudi Arabia. Infection & chemotherapy. 2014;46(1):30-4.
118. Naidoo P, du Toit E, Dunbar R, Lombard C, Caldwell J, Detjen A, et al. A Comparison of Multidrug-Resistant Tuberculosis Treatment Commencement Times in MDRTBPlus Line Probe Assay and Xpert (R) MTB/RIF-Based Algorithms in a Routine Operational Setting in Cape Town. Plos One. 2014;9(7):9.
119. Cox HS, Daniels JF, Muller O, Nicol MP, Cox V, van Cutsem G, et al., editors. Impact of decentralized care and the Xpert MTB/RIF test on rifampicin-resistant tuberculosis

treatment initiation in Khayelitsha, South Africa. Open Forum Infectious Diseases; 2015: Oxford University Press.
120. Skenders G, Holtz TH, Riekstina V, Leimane V. Implementation of the INNO-LiPA Rif.TB line-probe assay in rapid detection of multidrug-resistant tuberculosis in Latvia. International Journal of Tuberculosis and Lung Disease. 2011;15(11):1546-52.
121. Hanrahan CF, Dorman SE, Erasmus L, Koornhof H, Coetzee G, Golub JE. The impact of expanded testing for multidrug resistant tuberculosis using genotype [correction of geontype] MTBDRplus in South Africa: an observational cohort study. PLoS One. 2012;7(11):e49898.
122. Jacobson KR, Theron D, Kendall EA, Franke MF, Barnard M, van Helden PD, et al. Implementation of genotype MTBDRplus reduces time to multidrug-resistant tuberculosis therapy initiation in South Africa. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2013;56(4):503-8.
123. Kipiani M, Mirtskhulava V, Tukvadze N, Magee M, Blumberg HM, Kempker RR. Significant clinical impact of a rapid molecular diagnostic test (Genotype MTBDRplus assay) to detect multidrug-resistant tuberculosis. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2014;59(11):1559-66.
124. Singla N, Satyanarayana S, Sachdeva KS, Van den Bergh R, Reid T, Tayler-Smith K, et al. Impact of introducing the line probe assay on time to treatment initiation of MDR-TB in Delhi, India. PLoS One. 2014;9(7):e102989.
125. Lyu J, Kim MN, Song JW, Choi CM, Oh YM, Lee SD, et al. GenoType(R) MTBDRplus assay detection of drug-resistant tuberculosis in routine practice in Korea. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2013;17(1):120-4.
126. Raizada N, Sachdeva KS, Chauhan DS, Malhotra B, Reddy K, Dave PV, et al. A multi-site validation in India of the line probe assay for the rapid diagnosis of multi-drug resistant tuberculosis directly from sputum specimens. PLoS One. 2014;9(2):e88626.
127. Barnard M, Albert H, Coetzee G, O'Brien R, Bosman ME. Rapid molecular screening for multidrug-resistant tuberculosis in a high-volume public health laboratory in South Africa. American journal of respiratory and critical care medicine. 2008;177(7):787-92.
128. Anek-Vorapong R, Sinthuwattanawibool C, Podewils LJ, McCarthy K, Ngamlert K, Promsarin B, et al. Validation of the GenoType MTBDRplus assay for detection of MDR-TB in a public health laboratory in Thailand. BMC infectious diseases. 2010;10:123.
129. Tukvadze N, Kempker RR, Kalandadze I, Kurbatova E, Leonard MK, Apsindzelashvili R, et al. Use of a molecular diagnostic test in AFB smear positive tuberculosis suspects greatly reduces time to detection of multidrug resistant tuberculosis. PLoS One. 2012;7(2):e31563.
130. Seoudi N, Mitchell SL, Brown TJ, Dashti F, Amin AK, Drobniewski FA. Rapid molecular detection of tuberculosis and rifampicin drug resistance: retrospective analysis of a national U.K. molecular service over the last decade. Thorax. 2012;67(4):361-7.
131. Chryssanthou E, Angeby K. The GenoTypeReg. MTBDRplus assay for detection of drug resistance in Mycobacterium tuberculosis in Sweden. APMIS, Acta Pathologica, Microbiologica et Immunologica Scandinavica. 2012;120(5):405-9.
132. Gauthier M, Somoskovi A, Berland JL, Ocheretina O, Mabou MM, Boncy J, et al. Stepwise implementation of a new diagnostic algorithm for multidrug-resistant tuberculosis in Haiti. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2014;18(2):220-6.
133. Martinez-Lirola MJ, Munoz-Davila MJ, Garcia-De Viedma D, Cabezas Fernandez T, Luzon Garcia P. Usefulness of Genotype MTBDRplus assay in acid-fast bacilli positive

smear specimens in Almeria, Spain. Enfermedades Infecciosas y Microbiologia Clinica. 2014;32(8):511-4.
134. Singhal R, Myneedu VP, Arora J, Singh N, Sah GC, Sarin R. Detection of multi-drug resistance & characterization of mutations in Mycobacterium tuberculosis isolates from North- Eastern States of India using GenoType MTBDRplus assay. The Indian journal of medical research. 2014;140(4):501-6.
135. 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;377(9776):1495-505.
136. Hillemann D, Weizenegger M, Kubica T, Richter E, Niemann S. Use of the genotype MTBDR assay for rapid detection of rifampin and isoniazid resistance in Mycobacterium tuberculosis complex isolates. Journal of clinical microbiology. 2005;43(8):3699-703.
137. Organization WH, Organization WH. Improving the diagnosis and treatment of smear-negative pulmonary and extrapulmonary tuberculosis among adults and adolescents: recommendations for HIV-prevalent and resource-constrained settings. 2007.
138. Hutcheon JA, Chiolero A, Hanley JA. Random measurement error and regression dilution bias. Bmj. 2010;340.
139. Durovni B, Saraceni V, Cordeiro-Santos M, Cavalcante S, Soares E, Lourenço C, et al. Operational lessons drawn from pilot implementation of Xpert MTB/Rif in Brazil. Bulletin of the World Health Organization. 2014;92(8):613-7.
140. Organization WH. WHO monitoring of Xpert MTB/RIF roll-out. URL:< http://www who int/tb/laboratory/mtbrifrollout/en/index html. 2014.
141. Salje H, Andrews JR, Deo S, Satyanarayana S, Sun AY, Pai M, et al. The importance of implementation strategy in scaling up Xpert MTB/RIF for diagnosis of tuberculosis in the Indian health-care system: a transmission model. The Lancet infectious diseases. 2014;11(7):e1001674.
142. World Health Organization. Roadmap for rolling out Xpert MTB/RIF for rapid diagnosis of TB and MDR-TB.2010. Available from: http://www.who.int/tb/laboratory/roadmap_xpert_mtb_rif_rev23dec2010.pdf.
143. Blakemore R, Story E, Helb D, Kop J, Banada P, Owens MR, et al. Evaluation of the analytical performance of the Xpert MTB/RIF assay. Journal of clinical microbiology. 2010;48(7):2495-501.
144. 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;363(11):1005-15.
145. Steingart K, Sohn H, Schiller I, Kloda L, Boehme C, Pai M, et al. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. status and date: New, published in. 2013(1).
146. Food and Drug Administration. FDA permits marketing of first U.S. Test labeled for simultaneous detection of tuberculosis bacteria and resistance to the antibiotic rifampin. . 2013.
147. Armand S, Vanhuls P, Delcroix G, Courcol R, Lemaitre N. Comparison of the Xpert MTB/RIF test with an IS6110-TaqMan real-time PCR assay for direct detection of Mycobacterium tuberculosis in respiratory and nonrespiratory specimens. Journal of clinical microbiology. 2011;49(5):1772-6.
148. Moure R, Martín R, Alcaide F. Effectiveness of an Integrated Real-Time PCR Method for Detection of the Mycobacterium tuberculosis Complex in Smear-Negative

Extrapulmonary Samples in an Area of Low Tuberculosis Prevalence. Journal of clinical microbiology. 2012;50(2):513-5.
149. Pai M, Minion J, Jamieson F, Wolfe J, Behr M. Canadian Tuberculosis Standards 7th Edition. Canadian Resp Journal. 2013;20(Suppl A):16A-22A.
150. 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. Clinical Infectious Diseases. 2008;47(9):1135-42.
151. Clouse K, Page-Shipp L, Dansey H, Moatlhodi B, Scott L, Bassett J, et al. Implementation of Xpert MTB/RIF for routine point-of-care diagnosis of tuberculosis at the primary care level. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2012;102(10):805-7.
152. Cepheid. Xpert MTB/RIf assay [package insert]. Sunnyvale, CA2013.
153. World Health Organization. Policy Framework for Implementing New Tuberculosis Diagnostics. Geneva: WHO, 2010.
154. Kent PT, Kubica GP, Control CfD. Public health mycobacteriology: a guide for the level III laboratory: US Department of Health and Human Services, Public Health Service, Centers for Disease Control; 1985.
155. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. Ann Intern Med. 2003;138(1):40-4.
156. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. Ann Intern Med. 2003;138(1):W1-12.
157. Friedrich SO, Rachow A, Saathoff E, Singh K, Mangu CD, Dawson R, et al. Assessment of the sensitivity and specificity of Xpert MTB/RIF assay as an early sputum biomarker of response to tuberculosis treatment. The Lancet Respiratory Medicine. 2013;1(6):462-70.
158. Bark C, Gitta P, Ogwang S, Nsereko M, Thiel B, Boom W, et al. Comparison of time to positive and colony counting in an early bactericidal activity study of anti-tuberculosis treatment. The International Journal of Tuberculosis and Lung Disease. 2013;17(11):1448-51.
159. Zar HJ, Workman L, Isaacs W, Dheda K, Zemanay W, Nicol MP. Rapid diagnosis of pulmonary tuberculosis in African children in a primary care setting by use of Xpert MTB/RIF on respiratory specimens: a prospective study. The lancet global health. 2013;1(2):e97-e104.
160. Williamson D, Roberts S, Bower J, Vaughan R, Newton S, Lowe O, et al. Clinical failures associated with rpoB mutations in phenotypically occult multidrug-resistant Mycobacterium tuberculosis. The International Journal of Tuberculosis and Lung Disease. 2012;16(2):216-20.
161. Van Rie A, Mellet K, John M, Scott L, Page-Shipp L, Dansey H, et al. False-positive rifampicin resistance on Xpert® 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(2):206.
162. Van Deun A, Maug AK, Bola V, Lebeke R, Hossain MA, de Rijk WB, et al. Rifampicin drug resistance tests for tuberculosis: challenging the gold standard. J Clin Microbiol. 2013:JCM. 00553-13.
163. Somoskovi A, Deggim V, Ciardo D, Bloemberg GV. Inconsistent Results with the Xpert-MTB/Rif Assay in Detection of Mycobacterium tuberculosis with an rpoB Mutation Associated with Low Level of Rifampin Resistance: Diagnostic Implications. J Clin Microbiol. 2013:JCM. 01377-13.

164. Centers for Disease Control and Prevention. Availability of an assay for detecting Mycobacterium tuberculosis, including rifampin-resistant strains, and considerations for its use-United States, 2013. MMWR Morbidity and mortality weekly report. 2013;62(41):821.
165. Choi H, Miele K, Dowdy D, Shah M. Cost-effectiveness of Xpert® MTB/RIF for diagnosing pulmonary tuberculosis in the United States. The international journal of tuberculosis and lung disease: the official journal of the International Union against Tuberculosis and Lung Disease. 2013;17(10):1328.
166. Boehme CC, Nabeta P, Henostroza G, Raqib R, Rahim Z, Gerhardt M, et al. Operational feasibility of using loop-mediated isothermal amplification for diagnosis of pulmonary tuberculosis in microscopy centers of developing countries. Journal of clinical microbiology. 2007;45(6):1936-40.
167. Davis JL, Cattamanchi A, Cuevas LE, Hopewell PC, Steingart KR. Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis. The Lancet infectious diseases. 2013;13(2):147-54.
168. Qin ZZ, Pai M, Van Gemert W, Sahu S, Ghiasi M, Creswell J. How is Xpert MTB/RIF being implemented in 22 high tuberculosis burden countries? European Respiratory Journal. 2014:erj01477-2014.
169. Kik SV, Denkinger CM, Casenghi M, Vadnais C, Pai M. Tuberculosis diagnostics: which target product profiles should be prioritised? Eur Respir J. 2014;44(2):537-40.
170. Enosawa M, Kageyama S, Sawai K, Watanabe K, Notomi T, Onoe S, et al. Use of loop-mediated isothermal amplification of the IS900 sequence for rapid detection of cultured Mycobacterium avium subsp. paratuberculosis. Journal of clinical microbiology. 2003;41(9):4359-65.
171. Mori Y, Nagamine K, Tomita N, Notomi T. Detection of loop-mediated isothermal amplification reaction by turbidity derived from magnesium pyrophosphate formation. Biochemical and biophysical research communications. 2001;289(1):150-4.
172. Mitarai S, Okumura M, Toyota E, Yoshiyama T, Aono A, Sejimo A, et al. Evaluation of a simple loop-mediated isothermal amplification test kit for the diagnosis of tuberculosis. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2011;15(9):1211-7, i.
173. WHO. Guidelines for surveillance of drug resistance in tuberculosis (z.d.) 2011 [cited 2015 March 25]. Available from: http://www.who.int/tb/publications/2009/surveillance_guidelines/ en/index.html.
174. Shete PB, Farr, K., Strnad, L., Cattamanchi, A.,. A Systematic Review of TB-LAMP Diagnostic Accuracy for Pulmonary Tuberculosis. 2015.
175. World Health Organization. Tuberculosis country profiles 2014 [cited 2015 May 15]. Available from: http://www.who.int/tb/country/data/profiles/en/.
176. Sohn H, Puri, L.,. Cost and Affordability of TB-LAMP Assay. 2015.
177. Van Cleeff M, Kivihya-Ndugga L, Meme H, Odhiambo J, Klatser P. The role and performance of chest X-ray for the diagnosis of tuberculosis: a cost-effectiveness analysis in Nairobi, Kenya. BMC infectious diseases. 2005;5(1):111.
178. Laurence Y, Griffiths U, Vassall A. Costs to Health Services and the Patient of Treating Tuberculosis: A Systematic Literature Review. PharmacoEconomics. 2015;33(9):939-55.
179. Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, et al. Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. The Lancet. 2013;380(9859):2129-43.
180. UN operational exchange rate [Internet]. 2014 [cited May 10, 2015]. Available from: http://treasury.un.org/operationalrates/OperationalRates.aspx.

181. World Bank. The World Bank open data: indicator [cited 2015 December 18]. Available from: http://data.worldbank.org/indicator.
182. Pinto M, Steffen RE, Cobelens F, van den Hof S, Entriger A, Trajman A. Cost-effectiveness of the Xpert MTB/RIF assay for tuberculosis diagnosis in Brazil. International Journal of Tuberculosis and Lung Disease. 2015.
183. van't Hoog AH, Cobelens F, Vassall A, van Kampen S, Dorman SE, Alland D, et al. Optimal triage test characteristics to improve the cost-effectiveness of the Xpert MTB/RIF assay for TB diagnosis: a decision analysis. PLoS One. 2013;8(12):e82786.
184. Sohn H, Minion J, Albert H, dheda K, Pai M. TB diagnostic tests: how do we figure out their costs? . *Expert Rev Anti-Infect Ther*. 2009;7(6).
185. Cepheid. Xpert MTB/RIF Brochure [03/12/12]. Available from: http://www.cepheid.com/media/files/eu/brochures/XpertMTB_Broch_R9_EU.pdf.
186. van Soolingen D, Hermans PW, de Haas PE, Soll DR, van Embden JD. Occurrence and stability of insertion sequences in Mycobacterium tuberculosis complex strains: evaluation of an insertion sequence-dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J Clin Microbiol. 1991;29(11):2578-86.
187. Ahmad S, Araj GF, Akbar PK, Fares E, Chugh TD, Mustafa AS. Characterization of rpoB mutations in rifampin-resistant Mycobacterium tuberculosis isolates from the Middle East. Diagn Microbiol Infect Dis. 2000;38(4):227-32.
188. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007;23(21):2947-8.
189. Cepheid. Xpert MTB/RIF training brochure [03/12/12]. Available from: http://www.moleculartb.org/gb/pdf/ppt/14_SYMP_ISoundiram_Xperttraining_2902.pdf.
190. Lew W, Pai M, Oxlade O, Martin D, Menzies D. Initial drug resistance and tuberculosis treatment outcomes: systematic review and meta-analysis. Annals of internal medicine. 2008;149(2):123-34.
191. Nathanson E, Lambregts-van Weezenbeek C, Gupta R, Blöndal K, Caminero JA, Cegielski JP, et al. Multidrug-resistant tuberculosis management in resource-limited settings. 2006.
192. Menzies D, Benedetti A, Paydar A, Royce S, Pai M, Burman W, et al. Standardized treatment of active tuberculosis in patients with previous treatment and/or with mono-resistance to isoniazid: a systematic review and meta-analysis. PLoS Med. 2009;6(9):e1000150.
193. Espinal MA, Kim SJ, Suarez PG, Kam KM, Khomenko AG, Migliori GB, et al. Standard short-course chemotherapy for drug-resistant tuberculosis: treatment outcomes in 6 countries. Jama. 2000;283(19):2537-45.
194. Dye C, Garnett GP, Sleeman K, Williams BG. Prospects for worldwide tuberculosis control under the WHO DOTS strategy. The Lancet. 1998;352(9144):1886-91.
195. Dye C, Williams BG. Criteria for the control of drug-resistant tuberculosis. Proceedings of the National Academy of Sciences. 2000;97(14):8180-5.
196. Salomon JA, Lloyd-Smith JO, Getz WM, Resch S, Sánchez MS, Porco TC, et al. Prospects for advancing tuberculosis control efforts through novel therapies. PLoS Med. 2006;3(8):e273.
197. Abu-Raddad LJ, Sabatelli L, Achterberg JT, Sugimoto JD, Longini IM, Dye C, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. Proceedings of the National Academy of Sciences. 2009;106(33):13980-5.
198. Cleary SM, McIntyre D, Boulle AM. The cost-effectiveness of antiretroviral treatment in Khayelitsha, South Africa–a primary data analysis. Cost effectiveness and resource allocation. 2006;4(1):1.

Appendix

Appendix A. Full electronic search strategy

Appendix B. Graphic representation of total diagnostic and therapeutic delays for drug susceptible and drug resistant Tuberculosis

Appendix C: Set-up of testing in TB clinic

Appendix D: Sequencing on *rpoB* gene

Appendix E. Supplementary materials – Cost-effectiveness analysis of TB-LAMP

Appendix F. List of publication for Hojoon Sohn during his doctoral training period

Appendix A. Full electronic search strategy

Database: Embase Classic+Embase <1947 to 2015 January 29>

Search Strategy:

---

1 exp tuberculosis/ (232802)
2 Mycobacterium tuberculosis/ (57525)
3 tuberculosis.mp. (265363)
4 tb.tw. (58072)
5 or/1-4 (301560)
6 nucleic acid amplification/ (4872)
7 molecular diagnosis/ (4583)
8 nucleic acid test*.tw. (886)
9 NAAT.tw. (312)
10 NAATs.tw. (208)
11 NAA.tw. (6639)
12 direct amplification.tw. (244)
13 transcription-mediated amplification.tw. (341)
14 RNA amplification*.tw. (629)
15 DNA amplification*.tw. (3916)
16 molecular assay*.tw. (2074)
17 molecular diagnos*.tw. (10596)
18 polymerase chain reaction*.tw. (186836)
19 PCR.tw. (431859)
20 PCRs.tw. (2951)
21 Xpert.tw. (463)
22 GeneXpert.tw. (200)
23 cepheid.tw. (340)
24 "MTB/RIF".tw. (218)
25 cobas.tw. (2822)
26 TaqMan.tw. (13731)
27 AMTD*.tw. (91)
28 MTD.tw. (6672)
29 Gen-Probe.tw. (753)
30 ligase chain reaction*.tw. (536)
31 LCx.tw. (2337)
32 line probe assay*.tw. (625)
33 LPA.tw. (4609)
34 LPAs.tw. (146)
35 AMTD*.tw. (91)

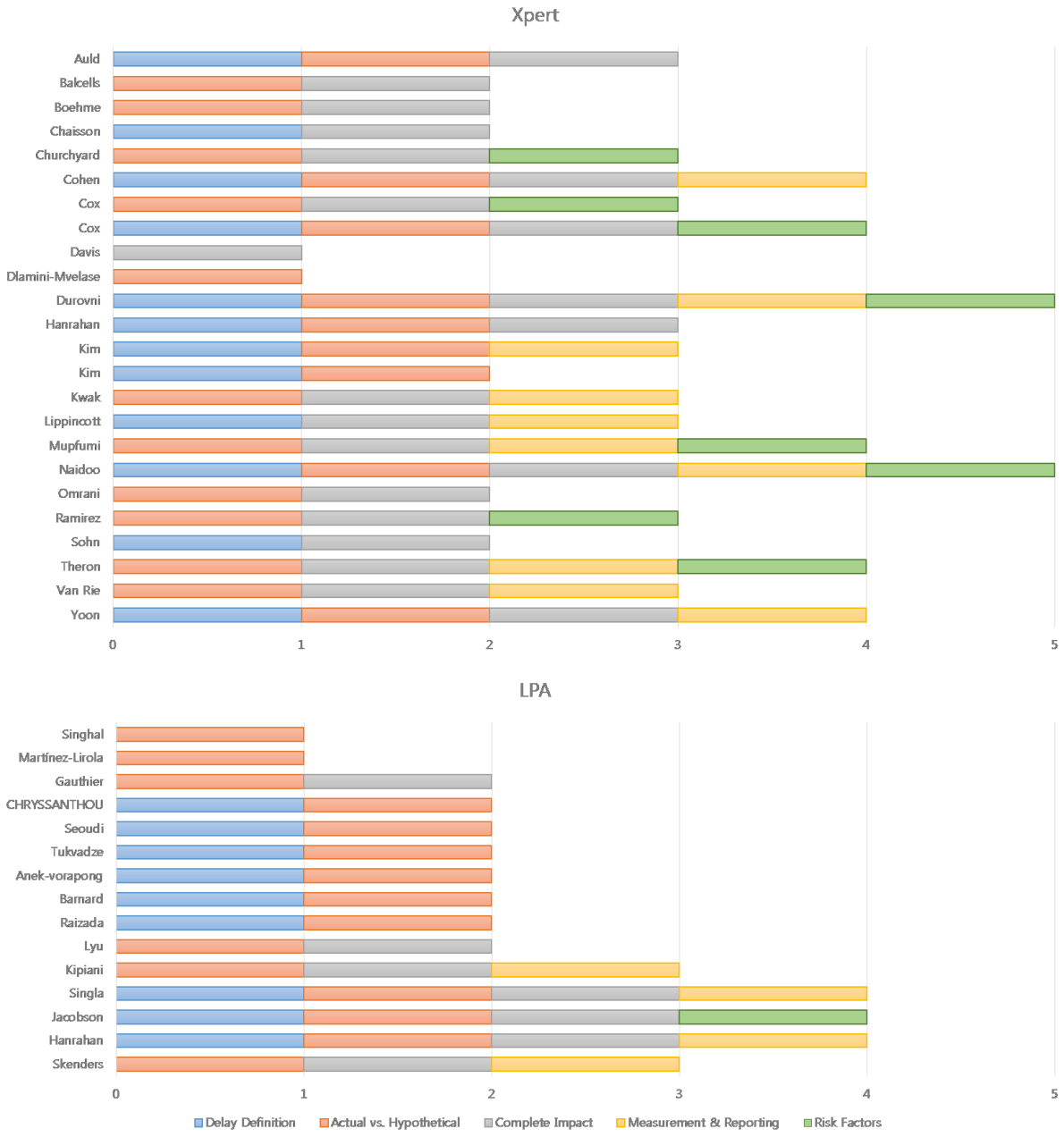
36 MTBDR*.tw. (176)
37 gMTBDR.tw. (1)
38 INNO-LiPA.tw. (580)
39 ProbeTec.tw. (117)
40 loopamp.tw. (12)
41 EXPAR.tw. (10)
42 LAMP.tw. (16871)
43 loop mediated amplification*.tw. (32)
44 Exponential Amplification Reaction*.tw. (9)
45 NALF.tw. (37)
46 nucleic acid lateral flow*.tw. (21)
47 (nucleic acid and amplification).tw. (5106)
48 (NAT or NATs).ti. (678)
49 (amplified and direct test*).tw. (142)
50 BD Probe.tw. (5)
51 Tec Direct.tw. (0)
52 or/6-51 (585569)
53 exp time/ (525789)
54 comparative effectiveness/ (8107)
55 exp "evaluation and follow-up"/ (1457223)
56 exp "treatment outcome"/ (979386)
57 exp morbidity/ (218643)
58 feasibility study/ (49946)
59 exp mortality/ (684720)
60 contact examination/ (2529)
61 exp infection control/ (80445)
62 cross infection/ (22558)
63 hospital information system/ (18032)
64 comparative study/ (743598)
65 intermethod comparison/ (179489)
66 exp survival/ (604552)
67 time to treatment/ (783)
68 exp diagnostic error/ (60382)
69 clinical decision making/ (16127)
70 medical decision making/ (65597)
71 decision making/ (140931)
72 exp "quality of life"/ (254950)
73 morbidity.tw. (327569)
74 feasibility.tw. (126601)

75 time.tw. (2635179)
76 mortality.tw. (637272)
77 outcome*.tw. (1189970)
78 conversion.tw. (164846)
79 follow-up.tw. (842102)
80 followup.tw. (32105)
81 decision*.tw. (270132)
82 impact.tw. (640283)
83 impacts.tw. (54478)
84 convert.tw. (24620)
85 delay*.tw. (437321)
86 adverse effect*.tw. (130415)
87 isolation.tw. (235688)
88 contact investigation*.tw. (572)
89 default.tw. (8909)
90 dropout*.tw. (8173)
91 drop-out*.tw. (6479)
92 empiric therapy.tw. (2227)
93 cure.tw. (94662)
94 failure*.tw. (695553)
95 relapse*.tw. (162955)
96 harm*.tw. (127568)
97 prevention.tw. (465774)
98 prevented.tw. (198315)
99 secondary case*.tw. (1603)
100 effectiveness.tw. (346458)
101 death*.tw. (738817)
102 undertreat*.tw. (5180)
103 under treat*.tw. (11600)
104 overtreat*.tw. (3330)
105 over treat*.tw. (1937)
106 adverse event*.tw. (116356)
107 adverse outcome*.tw. (18732)
108 undesirable effect*.tw. (3368)
109 patient centred.tw. (2978)
110 patient centered.tw. (6705)
111 contact tracing.tw. (1348)
112 contact examination*.tw. (140)
113 infection control.tw. (17400)

114 cross infection*.tw. (2817)
115 treatment fail*.tw. (26267)
116 recurrence.tw. (255869)
117 "point of care".tw. (8801)
118 survival.tw. (787636)
119 comparative stud*.tw. (102549)
120 "quality of life".tw. (206653)
121 qol.tw. (30987)
122 hrqol.tw. (10399)
123 or/53-122 (9669144)
124 5 and 52 and 123 (4402)
125 124 not (animal not human).sh. (4307)

****

Appendix B. Graphic representation of total diagnostic and therapeutic delays for drug susceptible and drug resistant Tuberculosis



A summary of the overall quality assessment score by each study
(maximum 5, minimum 0)

Appendix C. Set-up of testing in TB clinic

All of the sample and cartridge preparation was done in the negative-pressure room used for sputum induction (Figure S1 A). The patient was asked to provide a second induced sputum sample in a separate collection container. The sample preparation was done by the test operator in the sputum induction room after the patient left the room. The test operator was wearing a fitted N95 mask for personal protection. The operating procedure followed the recommendation for unprocessed sputum samples (185). The tuberculocidal sample reagent was added to the sample in a 2:1 ratio while still in the negative-pressure room. The sample was incubated for 15minutes. After incubation 2 ml of the sample was transferred to cartridge using sterile pipette. The cartridge was then taken out of the sputum induction room into a room without special ventilation (Figure S1 B), adjacent to the induction room within the TB clinic where the test was run on an open bench on a four-module GeneXpert machine.

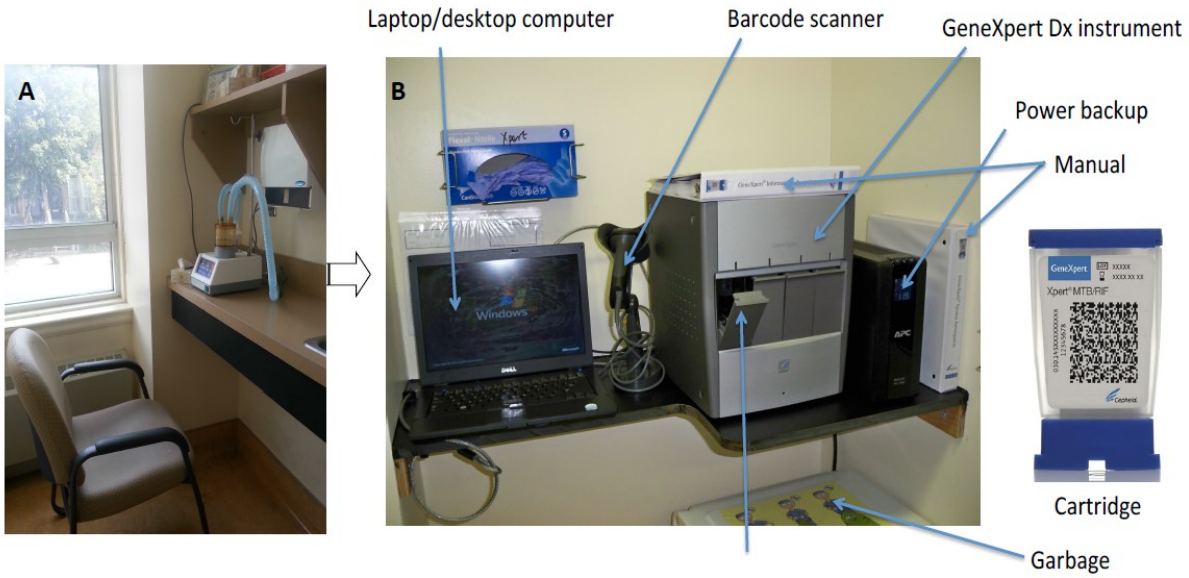


Figure E1: Sample collection and set up of GeneXpert in the routine TB clinic setting

Appendix D. Sequencing on rpoB gene

To determine the basis for the single isolate with discordant results from Xpert and DST rifampin resistance testing, we sequenced the 81 bp hypervariable region of the *rpoB* gene. Briefly, beginning with bacteria cultured in MGIT vials, DNA was extracted using the reference mycobacterial DNA extraction method (186). Next, the *rpoB* hypermutable region was amplified using oligonucleotide primers RPOBF: 59 AAACCAGATCCGGGTAGGCATG and RPO3R: 59 GTACGGCGTTTCGATGAACCCG, following the protocol described by Ahamd *et al*. (187). The 381 bp band was confirmed through 0.8% agarose gel electrophoresis and then sent for Sanger Sequencing at the McGill University and Genome Quebec Innovation Centre. The sequence was subsequently aligned to the H37rv reference genome (NC_000962.3) and the reference sequence from Xpert using ClustalW2 (188, 189). Based on these alignments, a C1534T mutation was identified, corresponding to a true-positive readout by probe A of Xpert (189). This mutation results in a Pro > Leu substitution at amino acid 511.

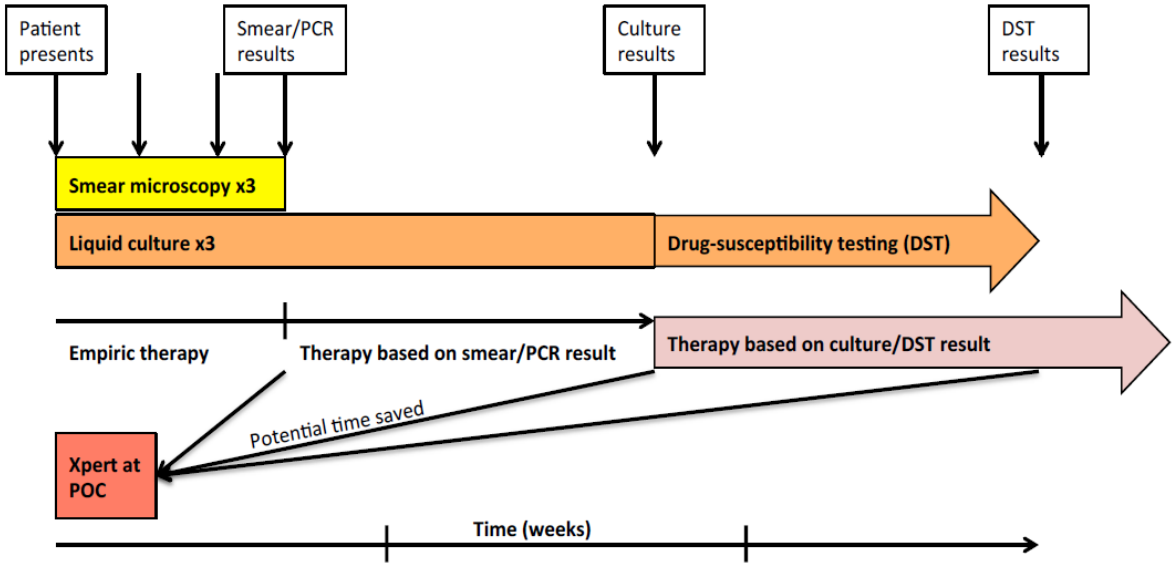
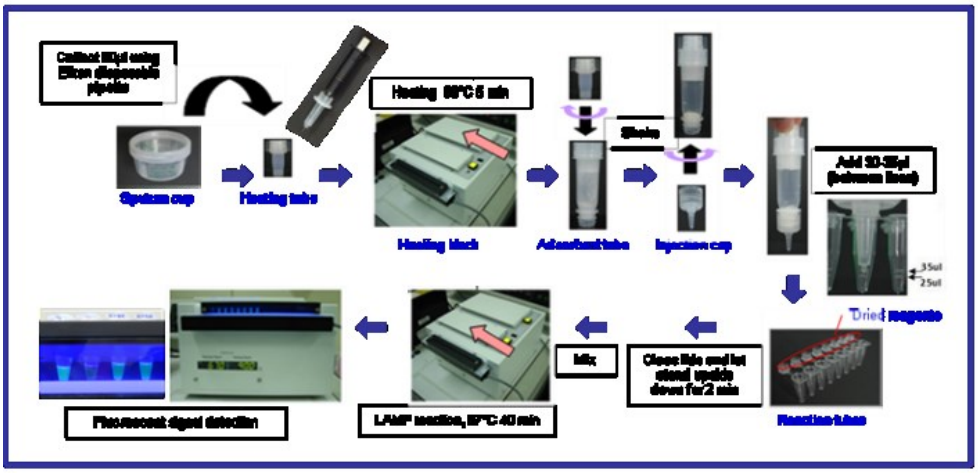


Figure E2: Time to diagnosis and treatment initiation

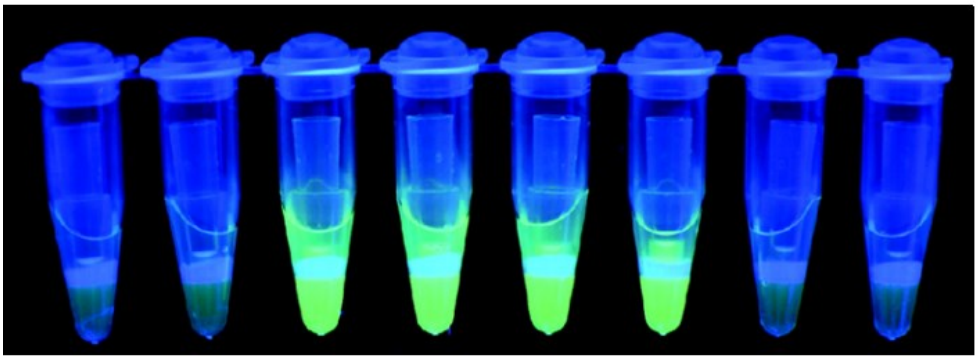
Outline of the different times it takes to obtain results from smear microscopy, culture and drug-susceptibility testing (DST). Treatment might be started empirically or based on smear- or PCR-result (if positive) or once culture result returns. Treatment might be modified based on a DST result. A positive Xpert result could potentially lead to a reduction of the time to initiation of treatment for drug-sensitive tuberculosis (TB) and in particular for drug-resistant TB if resistance is not suspected.

Appendix E. Supplementary materials – Cost-effectiveness analysis of TB-LAMP

A. Schematic diagram illustrating the LAMP procedural steps



B. Visual detection of LAMP product under UV light.
From left to right, tubes 1, 2, 7, and 8 are negative, and tubes 3, 4, 5, and 6 are positive



Reproduced from Boehme et al. JCM 2007

Figure A1. Illustration of the Loop-Mediated Isothermal Amplification (LAMP) technology and testing procedures

Type of TB	*Treatment regimen*	*Failure*		*Death distribution*		*Source*
		Probability (SE)	*Distribution*	*Probability (SE)*	*distribution*	
HIV-negative patients						
New drug-suspectible	first-line, category 1	**0.009 (0.001)**	beta	**0.029 (0.001)**	beta	(190)
	second-line	**0.050 (0.022)**	beta	**0.040 (0.019)**	beta	(191)
Previously treated drug-suspectible	first-line, category 2	**0.053 (0.009)**	beta	**0.048 (0.011)**	beta	(192)
	second-line	**0.078 (0.009)**	beta	**0.158 (0.013)**	beta	(191)
New drug-resistant	first-line, category 1	**0.277 (0.038)**	beta	**0.113 (0.027)**	beta	(193)
	second-line	**0.050 (0.022)**	beta	**0.040 (0.019)**	beta	(191)
Previously treated drug-resistant	first-line, category 2	**0.518 (0.048)**	beta	**0.164 (0.035)**	beta	(193)
	second-line	**0.078 (0.009)**	beta	**0.158 (0.013)**	beta	(191)
HIV-positive patients						
		probability (SE)	*distribution*	*probability (range)*	*distribution*	
New drug-suspectible	first-line, category 1	**0.009 (0.001)**	beta	**0.075 (0.045-0.087)**	uniform	(190)
	second-line	**0.050 (0.022)**	beta	**0.080 (0.040-0.120)**	triangular	(191)
Previously treated drug-suspectible	first-line, category 2	**0.053 (0.009)**	beta	**0.094 (0.050-0.140)**	triangular	(192) (193)
	second-line	**0.078 (0.009)**	beta	**0.204 (0.100-0.300)**	triangular	(191)
New drug-resistant	first-line, category 1	**0.277 (0.038)**	beta	**0.339 (0.226-0.452)**	triangular	(193)
	second-line	**0.050 (0.022)**	beta	**0.080 (0.040-0.120)**	triangular	(191)
Previously treated drug-resistant	first-line, category 2	**0.455 (0.047)**	beta	**0.339 (0.226-0.452)**	triangular	(193)
	second-line	**0.078 (0.009)**	beta	**0.316 (0.158-0.474)**	triangular	(191)

Table A. Treatment outcome probabilities used in the model.

1. Drug susceptible: no multidrug resistance, data based on treatment outcomes among pan-susceptible patients
2. Drug resistant: rifampicin resistance, assumed multidrug resistance; data based on treatment outcomes among multidrug resistant patients
3. Patients with drug-susceptible TB treated with second-line regimens are assumed to have similar failure and death rates as patients with MDR-TB, separately for new and previously treated cases
4. All treatment outcomes adjusted for defaulting and transfer out.
5. Failure rates for HIV-positive patients are assumed to be similar to those for HIV-negative patients, but adjusted downwards for previously treated drug-resistant cases treated with first-line category 2 regimen because distribution values for failure and cure would otherwise exceed 1.
6. Death rates for HIV-positive are derived by adding an excess death rate of 0.046 assuming ART (Akksilp 2007, Varma 2009, Abdool Karim 2010).

7. The mortality of HIV-infected MDR patients treated with first-line regimens is between 72 and 98% without antiretroviral treatment (Wells 2007). We assume antiretroviral treatment is given in conjunction with TB treatment, so this will lead to better survival. In studies from Thailand this resulted in about 5-fold reduction in mortality (Akksilp 2007, Varma 2009). Survival will however be less than when second-line treatment is given, which according to Seung et al. (2009) results in 2-fold increased mortality compared to HIV- patients. So we assume a 3-fold increase in mortality compared to HIV-negative patients (range 2-4). Similarly we assume a 2-fold increased death rate for HIV-infected MDR-TB patients treated with second-line regimens.

Patient category	Probability (range)	distribution
recovery, HIV-negative	0.10 (0.085-0.115)	Triangular
recovery, HIV-positive	0 (0-0.05)	Triangular
death, smear-positive TB, HIV-negative	0.25 (0.213-0.288)	Triangular
death, smear-positive TB, HIV-positive	1.0 (0.5-1.0)	Triangular
death, smear-negative TB, HIV-negative	0.10 (0.085-0.115)	Triangular
death, smear-negative TB, HIV-positive	0.67 (0.5-1.0)	Triangular

Table B. Probabilities used in the models of death and spontaneous recovery for undiagnosis (false-negative) untreated before a return diagnostic visit is made.

Source: Dye et al. Lancet 1998 (194); Dye & Williams PNAS 2000 (195); Salomon et al. PLoSMed 2006 (196); Abu-Raddad et al, PNAS 2009 (197).

Parameters	Vietnam	Malawi	Source
Age at onset (years)	37 (29, 48)	35 (28, 46)	WHO 2013
Life expectancy (years)	43.8	33.75	http://www.who.int/whosis/database/
Disability weight with TB, HIV-	0.331	0.271	(179)
Disability weight with TB, HIV+	0.399	0.135	(179)
Disability weight with AIDS	0.505	0.505	(179)
Disability weight HIV on ART	0.167	0.167	(179)
Median survival untreated HIV - Sm - (years)	7.36	7.36	(1)
Median survival untreated HIV - Sm + (years)	2.74	2.74	(1)
Median survival untreated HIV+ Sm- (years)	0.83	0.83	(1)
Median survival untreated HIV+ Sm + (years)	0.50	0.50	(1)
Survival Treated TB/MDR-TB HIV + with HAART (years)	12.9	12.9	(198)

Table C. Variables used in DALY calculations

Appendix F. List of publication for Hojoon Sohn during his doctoral training period

(September 2009 – April 2012, September 2014 – February 2016):

1 Theron, G., Peter, J., …, **Sohn, H**., Pai, M., Stein, D., Dheda. *Psychological distress and its relationship with non-adherence to TB treatment: a multicentre study.* BMC Infectious Diseases, 2015. 15(1): p. 1-12.

2 Alvarez, G. G., Van Dyk, D. D., Desjardins, M., Yasseen, A. S., Aaron, S. D., Cameron, D. W., ..., **Sohn, H**., & Pai, M. (2015). The feasibility, accuracy and impact of Xpert MTB/RIF testing in a remote Aboriginal community in Canada. *CHEST Journal*.

3 **Sohn, H**., Aero, A. D., Menzies, D., Behr, M., Schwartzman, K., Alvarez, G. G., ... & Denkinger, C. M. (2014). Xpert MTB/RIF testing in a low TB incidence, high-resource setting: limitations in accuracy and clinical impact. *Clinical infectious diseases*, ciu022.

4 Steingart, K. R., **Sohn, H.,** Schiller, I., Kloda, L. A., Boehme, C. C., Pai, M., & Dendukuri, N. (2013). Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *status and date: New, published in*, (1)

5 **Sohn H**, Pai M, Dendukuri N, Kloda LA, Boehme CC, Steingart KR. Xpert MTB/RIF test for detection of pulmonary tuberculosis and rifampicin resistance (Protocol). *Cochrane Database of Systematic Reviews* 2012, Issue 1. Art. No.: CD009593. DOI: 10.1002/14651858.CD009593.

6 Pang, Y., Li, Q., Ou, X., **Sohn, H.,** Zhang, Z., Li, J., ... & Zhao, Y. L. (2013). Cost-effectiveness comparison of genechip and conventional drug susceptibility test for detecting multidrug-resistant tuberculosis in china. *PloS one*, *8*(7), e69267.

7 Xia, H., Song, Y. Y., Zhao, B., Kam, K. M., O'Brien, R. J., Zhang, Z. Y., ..., **Sohn, H**., & Zhao, Y. L. (2013). Multicentre evaluation of Ziehl-Neelsen and light-emitting diode fluorescence microscopy in China. *The International Journal of Tuberculosis and Lung Disease*, *17*(1), 107-112.

8 Vassall A, van Kampen S, **Sohn H**, Michael JS, John KR, den Boon S, Davis JL, Whitelaw A, Nicol MP, Gler MT, Khaliqov A, Zamudio C, Perkins MD, Boehme CC, Cobelens F. Rapid diagnosis of tuberculosis with the Xpert MTB/RIF assay in high burden countries: a cost-effectiveness analysis.*PLoS Med.* 2011 Nov;8(11):e1001120. Epub 2011 Nov 8.

9 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.*EurRespir J.* 2011 Dec;38(6):1393-7. Epub 2011 Jun 9.

10 Pai M, Minion J, **Sohn H**, Zwerling A, Perkins M. Novel and improved technologies for TB diagnosis: Progress and Challenges. *Clin Chest Med.* 2009 Dec;30(4):701-16, viii.

11 **Sohn H**, Minion J, Albert H, Dheda K, Pai M. Tuberculosis Diagnostic Tests: How do we figure out their cost? *Exper Rev. Anti infect. Ther.* 7(6), 723-733 (2009)

12 Minion J, **Sohn H**, Pai M. Light-emitting diode technologies for TB diagnosis: what's on the market? *Exp Rev Med Devices* 2009;6(4):341-45.