Expanding the use cases of molecular testing for tuberculosis

Emily Lai Ho MacLean

Department of Epidemiology, Biostatistics and Occupational Health

Faculty of Medicine and Health Sciences

McGill University

March 2022

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

An estimated 10 million people developed tuberculosis (TB) in 2020, approximately 4 million of whom never received a diagnosis or notification. This may be partly attributed to continued reliance on conventional, inaccurate tests such as smear microscopy. While the World Health Organization (WHO) has approved several molecular TB diagnostic tests, their uptake is still limited. Molecular tests can quickly and accurately detect adult pulmonary TB, the most common manifestation of TB. However, improvement is needed in test uptake with regards to broadening the use cases for which molecular TB tests are deployed. Part of the issue is a lack of evidence for how these molecular tests perform in diagnosing conditions other than adult pulmonary TB, and how to optimize the integration of molecular TB testing with other infectious diseases like COVID-19. Therefore, the overall aim of this thesis was to understand the performance characteristics of molecular TB diagnostics when deployed in a variety of use cases and populations.

The first manuscript is a 2020 narrative review of molecular diagnostics for TB. The state of molecular diagnostics is described, including their performance characteristics, as well as key emerging technologies and themes.

In the second manuscript, we systematically reviewed the literature of studies reporting the diagnostic accuracy of Xpert MTB/RIF (Xpert), an automated PCR-based TB test, in detecting pediatric TB using stool samples. Xpert was designed to run on sputum, but young children often cannot spontaneously produce this sample. However, children swallow sputum while sleeping, so TB bacteria are detectable in stool. In a meta-analysis of results from 9 studies we found that stool-Xpert had a sensitivity of 67% and specificity of 99% against a microbiological reference standard. Stool-Xpert has most obvious utility as a rule-in test.

In the third manuscript, we turned to extrapulmonary TB (EPTB), for which there is no gold standard test. Thus, a composite reference standard (CRS) of multiple tests is typically deployed for diagnosis. Existing TB tests are not optimized for EPTB and, correspondingly, might perform poorly. This has implications for evaluating molecular tests like Xpert, as comparisons with an imperfect CRS will bias test accuracy estimates. We illustrated this issue by comparing CRS-based test accuracy estimates with estimates from Bayesian latent class analysis (LCA) using a dataset from Delhi, India. We constructed a heuristic diagram to distinguish target conditions of each test, then performed Bayesian LCA to reflect this model. LCA considered imperfect EPTB test accuracy and conditional dependence to estimate tests' sensitivities and specificities, while CRSs ignored these details and were difficult to interpret.

In the fourth manuscript, we investigated the integration of TB and COVID-19 testing with colleagues at the Universidad Peruana Cayetano Heredia in Lima, Peru. As many low- and middle-income countries now have a high burden of both TB and COVID-19, it is critical to understand how to integrate and optimize dual testing. With a CIHR/IDRC grant, our team

evaluated the combined use of Xpert MTB/RIF Ultra, a TB test, and Xpress SARS-CoV-2, a COVID-19 test, which are run on a common testing platform. We estimated the diagnostic accuracy of Xpress on sputum and assessed the diagnostic yield of using a single sputum sample to diagnose TB and COVID-19. Leveraging existing infrastructure to concurrently test for TB and COVID-19 may help maintain TB care.

Improved uptake of newer diagnostics will be needed to find the missing TB cases and meet the WHO End TB Goals. Molecular test deployment in various populations, such as children and people with EPTB, and in innovative interventions, such as multi-disease testing, is necessary to maintain TB care in the on-going COVID-19 pandemic. My thesis findings are forming part of the evidence base for expanding the use cases of molecular TB testing.

Resumé

Élargissement des types de cas pour l'utilisation de tests moléculaires pour la tuberculose

On estime que 10 millions d'individus ont développé la tuberculose (TB) en 2020, dont environ 4 millions n'ont pas été diagnostiqués. Ceci peut être partiellement dû au maintien de l'utilisation de tests conventionnels et imprécis, tels que la microscopie par frottis. Bien que l'Organisation mondiale de la santé (OMS) ait approuvé plusieurs tests diagnostiques moléculaires pour la TB, leur adoption demeure limitée. Les tests moléculaires peuvent détecter rapidement et de façon fiable la TB pulmonaire chez les adultes, la manifestation la plus répandue de la TB. Toutefois, l'adoption de ces tests doit être améliorée en ce qui a trait à l'élargissement des types de cas pour lesquels ces tests sont employés. Une partie du problème est le manque de données probantes démontrant comment ces tests moléculaires performent lors du diagnostic de conditions autres que la TB pulmonaire chez les adultes, et comment optimiser l'intégration de ces tests avec d'autres maladies infectieuses telles que la COVID-19. Ainsi, l'objectif principal de cette thèse était de comprendre les caractéristiques de performance des tests diagnostiques moléculaires pour la TB lorsqu'utilisés dans divers types de cas et de populations.

Le premier article est une revue narrative des tests diagnostiques moléculaires pour la TB publiée en 2020. Le portrait actuel de ces tests est décrit, incluant leurs caractéristiques de performance, ainsi que les technologies émergentes et thèmes clés.

Le second article est une revue de littérature systématique d'études rapportant la précision diagnostique de Xpert MTB/RIF (Xpert), un test PCR automatisé pour la TB, pour la détection de TB pédiatrique dans des échantillons de selles. Xpert a été conçu pour des échantillons de crachats, mais les jeunes enfants ne peuvent pas produire spontanément ce type d'échantillon. Toutefois, les enfants avalent des crachats durant leur sommeil, et la bactérie de TB est ainsi détectable dans les selles. Dans une méta-analyse des résultats de 9 études, nous avons trouvé que Xpert-selles avait une sensibilité de 67% et une spécificité de 99% comparativement à un test de référence microbiologique. Xpert-selles a donc une utilité claire comme test d'inclusion (peu de faux positifs).

Dans le troisième article, nous nous sommes penchés sur la TB extrapulmonaire (TBEP), pour laquelle il n'existe pas de test de référence. Dans ce cas, un test de référence composite de plusieurs tests est généralement utilisé pour le diagnostic. Les tests existants ne sont pas optimisés pour la TBEP et par le fait même pourraient mal performer. Ceci a des répercussions sur l'évaluation de tests moléculaires comme Xpert puisque la comparaison avec un test de référence composite imparfait biaisera l'estimation de la précision du test. Nous avons illustré ce problème en comparant des estimations de précision de tests basées sur des tests de référence composite avec des estimations par analyse bayésienne de classe latente utilisant un ensemble de données de Delhi, Inde. Nous avons construit un diagramme heuristique pour identifier les conditions cibles de chaque test, pour ensuite exécuter une analyse bayésienne de

classe latente afin de refléter ce modèle. L'analyse a pris en considération la précision imparfaite du test pour la TBEP et la dépendance conditionnelle pour estimer les performances des tests, alors que les tests de référence composite ignoraient ces détails et étaient difficiles à interpréter.

Dans le quatrième article, nous avons examiné l'intégration de tests pour la TB et la COVID-19 avec des collègues de l'Universidad Peruana Cayetano Heredia (Lima, Pérou). Puisque plusieurs pays à revenu faible ou intermédiaire ont maintenant à la fois une charge de morbidité importante de TB et de COVID-19, il est impératif de déterminer comment intégrer et optimiser le diagnostic combiné de ces deux maladies. À l'aide d'une subvention de CRDI/IRSC, notre équipe a évalué l'utilisation combinée de Xpert MTB/RIF Ultra, un test pour la TB, et Xpress SARS-CoV-2, un test pour la COVID-19, qui utilisent la même plateforme diagnostique. Nous avons estimé la précision diagnostique de Xpress avec des crachats et évalué le rendement diagnostique lors de l'utilisation d'un échantillon de crachat unique pour diagnostiquer à la fois la TB et la COVID-19. Tirer partir des infrastructures existantes pour tester en concomitance la TB et la COVID-19 sera critique pour maintenir les soins en TB.

L'amélioration de l'adoption de nouveaux outils diagnostiques sera nécessaire pour identifier les cas de TB non-diagnostiqués et atteindre les objectifs de l'OSM pour éradiquer la TB. Le déploiement des tests moléculaires dans diverses populations, comme les enfants et les individus atteints de TBEP, et dans des interventions innovatrices, telles que les tests multi-maladies, est nécessaire pour maintenir les soins en TB durant la pandémie de COVID-19 en cours. Les résultats de ma thèse contribuent au bassin de données probantes pour l'élargissement des types de cas pour l'utilisation de tests moléculaires pour la TB.

Acknowledgements

I cannot imagine a better graduate school experience than the one I'm about to complete. Naturally, this is due to all characters involved.

I first want to thank my supervisor, Dr Madhu Pai, for being a great example. Madhu shows that you can always do the right thing. I've learned from him that research can be rigorous while also compassionate. Madhu encourages everyone around him to do better – in the sense of productivity but, much more importantly, with respect to creating positive impact. I'm also extremely grateful for the many opportunities he presented me, the introductions to collaborators, and the financial support.

I'd also like to thank my co-supervisor, Dr Nandini Dendukuri. She has taught me so much through our discussions and endless emails. I cannot thank her enough for her encouragement, expertise, and patience. She has been instrumental in my development as a scientific thinker, and I can only aspire to replicate her vibe of extreme competence and serenity in my career.

As well, I must thank the best committee member ever, Dr Claudia Denkinger. I'm so grateful she would always make time to mentor and support me. She is an excellent role model and has taught me a lot about thinking critically and seeing the big picture. I value her exceptionally insightful feedback and aim to emulate the realism and humour she brings to problem-solving.

Many thanks also to past and present members of Team Pai! Working with them has been a true privilege and they impress and amuse me endlessly. Gratitude is particularly owed to Caroline Vadnais for keeping everything running and making all things doable.

Thank you to my fellow students and cohort colleagues. Their generosity, care, and reality checks have helped me through all the most challenging parts of the program.

To the team at UPCH, especially César, Luz, and Tati, I am thankful x1000 for everything. I have learned so much through working with them and I am inspired by their dedication. Even though our collaboration was remote, I never felt we were disconnected. Their perseverance, confidence, and adaptability in the face of a huge disease burden were truly instructive.

Thanks to my family for their love. I want to thank my parents, Alan and Peggy, for raising me to be a curious person and an independent thinker. I'm thankful for their blind confidence in me (even if they don't fully know what I'm up to), and for rarely making me feel like I needed to hurry up and graduate. I also want to thank Rachel and Madeline, my sisters, for always gassing me up, commiserating with me, and generally always being in my corner every day. I love you all.

I'm truly grateful to Dave for always listening keenly, noticing everything, and believing in me.

I don't know how to thank my wonderful, beguiling, and resolute friends. It's pretty easy to face anything a PhD throws at you and keep pushing yourself day-in-day-out, for years, when you're surrounded by people this brilliant. More long dinners, terrasse hangs, bike rides, late nights, and conversations in the future.

Financial support

I am extremely grateful for the financial support I received throughout my PhD. I held a McGill University Graduate Excellence Fellowship from 2016-2017 and a Harrison-Watson Faculty of Medicine Studentship from 2017-2018. From 2018 to end of 2021 I was supported by a Doctoral Fellowship from the Fonds de recherche du Québec – Santé. Throughout my degree, I received supplemental funding from my supervisor, Dr Madhu Pai. Additionally, I was supported by a Global Health Grads fellowship and multiple travel awards from the McGill International TB Centre for research and conference travel.

Preface and Contribution of authors

As first author on all manuscripts included in this thesis, I developed the research questions and protocols for all chapters, with feedback and support from my co-supervisors, Dr Madhu Pai and Dr Nandini Dendukuri, committee member Dr Claudia Denkinger, and manuscript co-authors. I also oversaw and managed the primary research study in Lima. All analyses were conducted by me. I was responsible for interpretation of the results, drafting of the manuscripts, and subsequent submissions and revisions for all published and submitted papers. All manuscripts included in this thesis were written by me. Detailed author contributions for each manuscript are described below.

Manuscript 1: Advances in molecular diagnosis of tuberculosis

This narrative mini-review was commissioned by the *Journal of Clinical Microbiology*. Dr Madhu Pai and I conceptualised the study and decided on content and themes. I coordinated the review and co-author contributions. I composed the first draft, with written contribution from Dr Anita Suresh in the next generation sequencing section and Dr Stefan F Weber in the concluding remarks. All authors provided critical feedback.

Manuscript 2: Diagnostic accuracy of stool Xpert MTB/RIF for the detection of pulmonary tuberculosis in children: a systematic review and meta-analysis

This study was designed by coauthors Dr Giorgia Sulis and Dr Faiz Ahmad-Khan, with input from me. I coordinated the project. Dr Sulis and I wrote the study protocol, developed the data collection tools, and executed the search strategy. We were the primary reviewers for publication screening, data extraction, and quality assessment. I cleaned the data, performed the meta-analysis, and drafted the manuscript. All authors provided critical feedback.

Manuscript 3: Bayesian latent class analysis versus composite reference standards for extra-pulmonary tuberculosis diagnostic tests

This study was conceived of and designed by Dr Dendukuri and me. Dr Dendukuri and I set the study objectives. Secondary data were provided by coauthors Dr Mikashmi Kohli and Dr Surendra K. Sharma. I performed all analyses, with programming support from Mr Ian Schiller when necessary. I interpreted the data, prepared figures and tables, and wrote the manuscript. Dr Sharma and Dr Claudia Denkinger provided clinical context and interpretation. All authors provided critical feedback.

Manuscript 4: Integration of molecular testing for M. tuberculosis and SARS-CoV-2 in Lima, Peru

I was the primary grant writer for the CIHR project grant that funded this study. The study protocol and design were established by me, with support from Dr Madhu Pai, Dr César Ugarte-Gil, and Dr Luz Villa. The study objectives were established by Dr Pai, Dr Mikashmi Kohli, Dr Giorgia Sulis, and me. I remotely managed the study (as restrictions due to the COVID-19

pandemic precluded travel), with Dr Ugarte-Gil and Dr Villa overseeing the patient recruitment and data collection. Dr Tatiana Caceres-Nakiche coordinated all laboratory testing. Dr Nandini Dendukuri and I designed the analysis which I conducted. All authors provided critical feedback.

Statement of originality

The four manuscripts that comprise the main body of this thesis are all original scholarship and provide contributions to knowledge.

My narrative review (Chapter 2, Manuscript 1) was a commissioned mini-review that aimed to synthesise the most recent developments in the field of molecular diagnostics for tuberculosis (TB).

My systematic review (Chapter 3, Manuscript 2) was the first study to meta-analyse the diagnostic accuracy of Xpert MTB/RIF using stool samples for diagnosing childhood TB. This manuscript, which pooled data from nine different studies, was important to informing the evidence base around alternative specimens for childhood TB diagnosis. Indeed, the data were later shared with the group performing a World Health Organisation-commissioned review on this topic.

The Bayesian latent class analysis (LCA) (Chapter 4, Manuscript 3) is the first to utilize this statistical modeling technique to determine the diagnostic accuracy of multiple tests for diagnosing extrapulmonary TB. After developing a heuristic model to clearly demonstrate our model's assumptions, we demonstrated the utility of Bayesian LCA in the absence of a gold standard for estimating multiple diagnostic tests' accuracies. We contrasted the results with those produced by composite reference standards.

My primary research study (Chapter 5, Manuscript 4) was a CIHR/IDRC-funded study that investigated integrating molecular testing for TB and COVID-19 with a single respiratory sample using a multiplex testing platform. Our study was one of the largest of its kind, and it was also the largest to assess the diagnostic accuracy of a new COVID-19 test, Xpress SARS-CoV-2, on sputum.

List of abbreviations

ATT – anti-tuberculosis therapy

CI – confidence interval

CrI – credible interval

COVID-19 – Coronavirus disease 2019

CRS – composite reference standard

CSF – cerebral spinal fluid

CXR – chest x-ray

CrI – credible interval

DR-TB – drug resistant TB

DST – drug susceptibility testing

INH – isoniazid

LAM – lipoarabinomannan

LCA – latent class analysis

LPA – line probe assay

LTBI – latent TB infection

MCMC – Markov chain Monte Carlo

MDR-TB – multi drug resistant tuberculosis

NA – not applicable

NAAT – nucleic acid amplification test

NE – not estimable

NGS – next generation sequencing

No. – number of

NTP – National TB Program

POC – point-of-care

PTB – pulmonary TB

QUADAS-2 – The Quality Assessment of Diagnostic Accuracy Studies-2

RIF – rifampicin

SLID – second line injectable drug

Smear – smear microscopy

TB – tuberculosis

TPP – target product profile

Ultra – Xpert MTB/RIF Ultra

WGS – whole genome sequencing

WHO – World Health Organization

XDR-TB – extensively drug resistant tuberculosis

Xpert – Xpert MTB/RIF

List of tables

List of figures

1 CHAPTER 1: INTRODUCTION

1.1 Introduction

Tuberculosis (TB) is one of the most significant global public health threats, and, though it does not receive much attention in our Western context, TB continues to rage on it many parts of the world. This may change, as the public has a new-found interest in infectious diseases. For now, though, TB's devastating impacts on patients and communities have worsened since the advent of the coronavirus disease 2019 (COVID-19) pandemic (1).

There are approximately 10 million incident cases of TB each year, mostly pulmonary forms in adults, and it is estimated that at least one third of these cases are never officially reported to National TB Programs (NTPs) (2). In these instances, patients are unlikely to receive proper diagnosis and appropriate care. This leaves untreated people with TB at high risk of morbidity and death (3) and allows transmission of TB to continue unabated (4).

A disease that has been with humans since prehistory (5), TB remains challenging to detect in many settings and populations. Diagnoses of TB in children or of extrapulmonary disease are especially uncertain because there are no perfect tests for these disease conditions (6, 7). In the last ten years, molecular diagnostic tests for TB have emerged, many of which have improved characteristics over traditional bacteriologic methods (8). Molecular tests have the potential to improve the quality-of-care people with TB receive, but the use cases for which these tools are deployed must be expanded (9).

1.2 Research gaps addressed by this thesis

There has been a flurry of activity in the molecular TB diagnostics space in the last decade. However, these tests were developed primarily to be used on sputum samples for the diagnosis of adult pulmonary TB (PTB). Unfortunately, children also suffer from TB, and it is often challenging for them to produce the requisite sputum samples for TB testing. TB also does not exclusively impact the lungs, and extrapulmonary disease forms are not rare. Therefore, other specimens have clinical relevance. As the COVID-19 pandemic continues to impact healthcare systems, molecular TB testing must continue, even as resources are diverted to the newer pandemic. During the COVID-19 pandemic, countries have greatly increased their capacities to conduct molecular testing, with millions of COVID-19 tests run each day. The TB community must learn from this experience and leverage the molecular capacity that now exists worldwide.

The accuracy of molecular tests using alternate specimens for childhood and extrapulmonary TB is not particularly well-characterised, particularly when considering that no perfect reference standard exists for these conditions. Additionally, the systematic integration of TB and COVID-19 molecular testing has not been investigated, although the diseases share common risk factors and symptoms.

1.3 Thesis goal and objectives

The overall goal of my thesis is to address certain gaps in the evidence base surrounding the use of molecular TB diagnostics for conditions other than adult pulmonary TB. In this manuscript-based thesis, I will address four objectives:

Objective 1 – Manuscript 1

To summarise the most recent advances in the field of molecular diagnostics for TB with respect to new and emerging assays, as well as contemporary themes in deployment of molecular TB testing

Objective 2 – Manuscript 2

To systematically review and meta-analyse the literature regarding the use of Xpert MTB/RIF on an alternative specimen type, namely, stool, for the detection of TB in children

Objective 3 – Manuscript 3

To utilize Bayesian latent class analysis to estimate the diagnostic accuracy of multiple tests, including Xpert MTB/RIF, for the detection of extrapulmonary TB in the absence of a perfect reference standard

Objective 4 – Manuscript 4

To estimate the diagnostic accuracy of a molecular COVID-19 test on an alternative specimen type, namely, sputum, and to investigate the integration of this COVID-19 test with molecular TB testing in Lima, Peru

2 CHAPTER 2: LITERATURE REVIEW

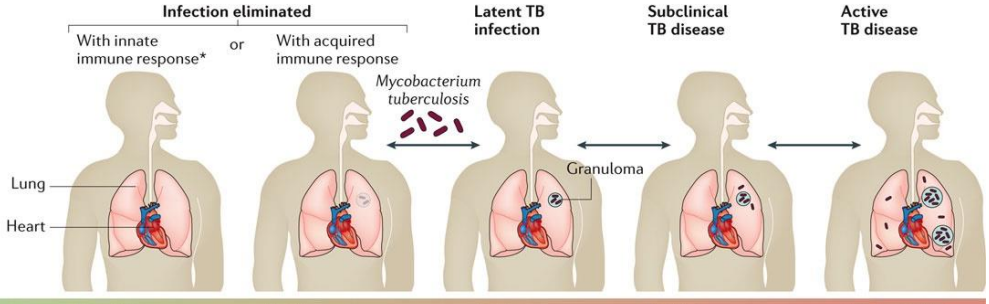
2.1 Tuberculosis – a global health problem

2.1.1 Background

Tuberculosis (TB) is an infectious disease caused by the bacteria *Mycobacterium tuberculosis* (*M. tuberculosis*). An airborne infection (10), transmission is thought to occur through the inhalation of aerosolized *M. tuberculosis*-containing droplets produced by an individual with TB (11). TB is most frequently a disease of the lungs, but extrapulmonary disease forms are also possible. Evidence accumulating since the beginning of the antibiotic era indicates that infected individuals who have not been diagnosed and thus have not started anti-TB therapy (ATT) are responsible for the majority of *M. tuberculosis* transmission (4, 12-15). Once enrolled on the appropriate ATT, people with TB quickly become less infectious (10, 16, 17).

Previously, it was thought that TB could only be transmitted by symptomatic people with active TB disease. However, it is increasingly understood that TB exists on a continuous disease spectrum (18). At one end is latent TB infection (LTBI) and active TB at the other (Figure 2.1), but a subclinical stage in between also intermittently contributes to disease spread (18, 19).

Therefore, in order to improve TB control efforts and decrease the global TB burden, cases must be quickly identified and treated. To that end, a key component of the World Health Organization's (WHO) End TB Strategy is improved roll-out and uptake of new tests for TB (20).



TST	Negative	Positive	Positive	Positive	Usually positive
IGRA	Negative	Positive	Positive	Positive	Usually positive
Culture	Negative	Negative	Negative	Intermittently positive	Positive
Sputum smear	Negative	Negative	Negative	Usually negative	Positive or negative
Infectious	No	No	No	Sporadically	Yes
Symptoms	None	None	None	Mild or none	Mild to severe
Preferred treatment	None	None	Preventive therapy	Multidrug therapy	Multidrug therapy

Figure 2.1: The spectrum of TB. TB exists on a continuum from *M. tuberculosis* infection to active (pulmonary) TB disease (21).

2.1.2 Global epidemiology of TB

As mentioned above, there were an estimated 10 million incident cases of TB in 2020, although only 5.8 million were officially notified to NTPs. This burden is not shared equally around the world (Figure 2.2), as cases in the 30 high TB burden countries (22) constitute 86% of the global total, and eight of those countries experienced two thirds of the global TB burden (2). Most TB occurs in adults, but children comprise approximately 16% of all cases. Extrapulmonary TB prevalence varies between regions, from 8% in the WHO Western Pacific region to 24% in the WHO Eastern Mediterranean region (23).

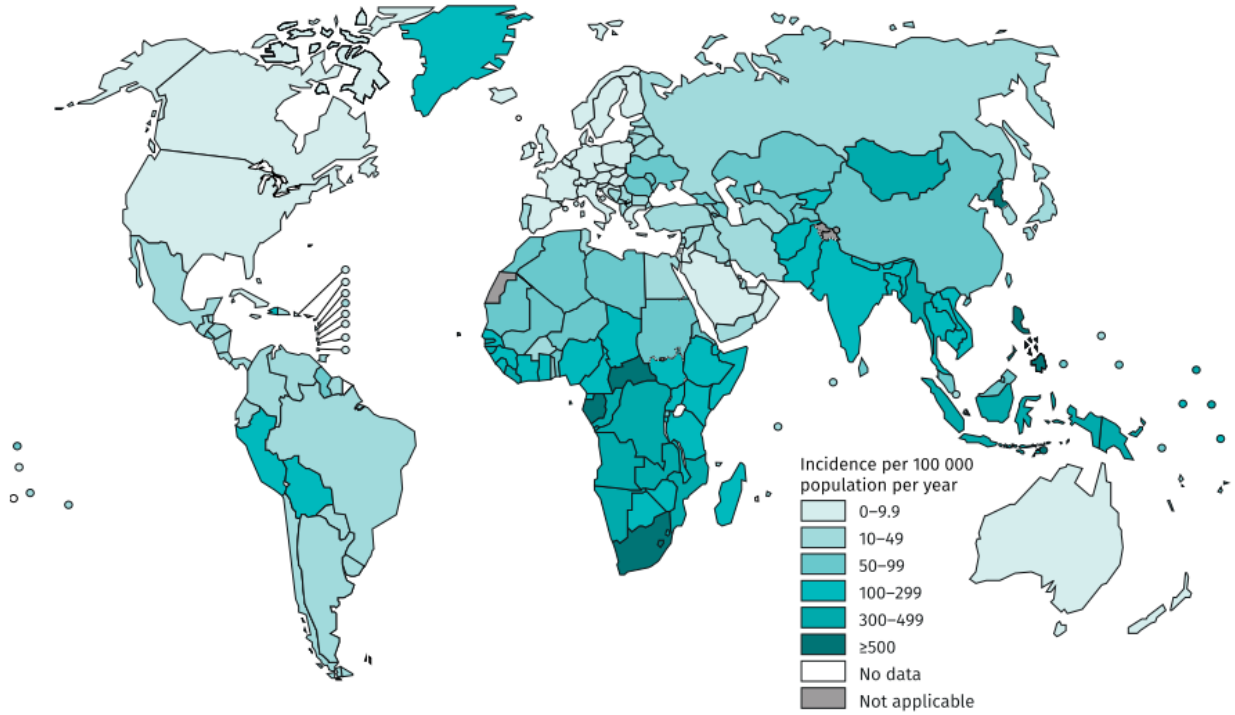


Figure 2.2: Estimated 2020 TB incidence rates by country (2).

Before the COVID-19 pandemic emerged, there were already an estimated 3 million incident cases of TB that were 'missing' (i.e., cases that arose in people who did not receive a diagnosis or were not notified to NTPs) each year.

Unfortunately, this gap in incident versus reported cases has increased significantly due to the pandemic, with large sudden drops in case notifications observed (Figure 2.3). The reduction in people accessing TB services and care has led to an increase in TB-related mortality to levels not seen since 2017: over 1.5 million people died of TB, a curable disease, in 2020.

Figure 2.3: Global trends in TB case notifications. In the top panels, case notifications from January 2020 to June 2021 for select high TB burden countries are shown in green, with the bold black line indicating the average number of monthly or quarterly case notifications in 2019. The bottom image shows the overall annual global trend in case notifications of people newly diagnosed with TB from 2016-2020 (2).

Cascade of care analyses, which evaluate patient retention across sequential stages of disease care, have been constructed for different countries to investigate where people with TB are being lost. A generic TB care cascade is pictured in Figure 2.4. Each step contains a total number of people at a particular phase in care, and a proportion of these individuals is lost in the between-step gaps (24). Molecular testing has the potential to mitigate losses in gaps 1 and 2; the development and roll-out of tests that can be run in peripheral settings of healthcare systems could help people in step 1 reach step 2, and increased use of highly accurate molecular tests would help more people in step 2 reach step 3. Closing gaps 1 and 2 could

substantially improve the proportion of people reaching recurrence-free survival. A care cascade analysis of the TB epidemic in India estimated that almost 30% of all people with TB never access testing, and a further 10% are lost due to never receiving a diagnosis (Figure 2.4) (25).

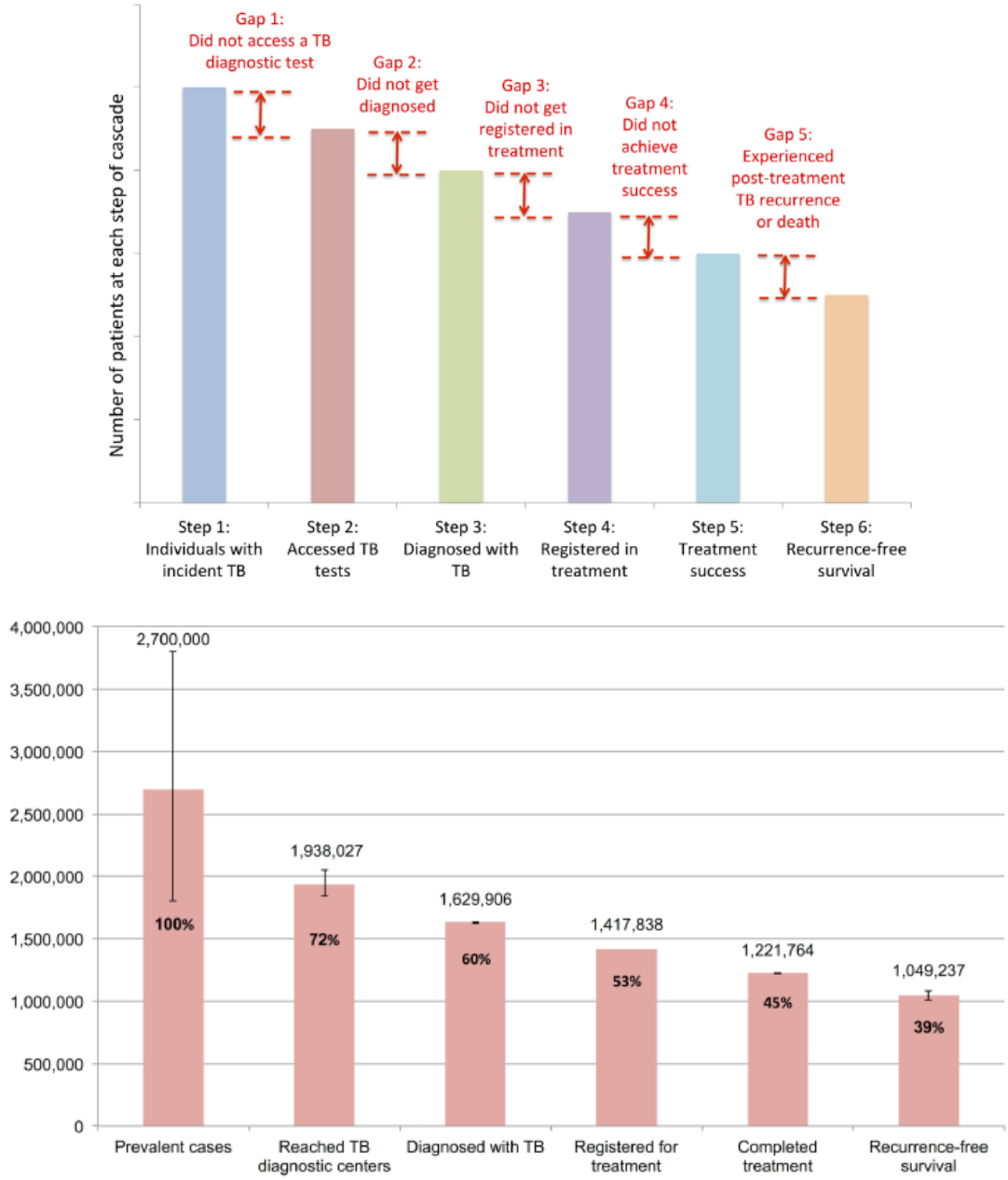


Figure 2.4: TB cascades of care. A generic care cascade (24) (top) and the TB care cascade in India's public healthcare sector (bottom) (25).

2.1.3 TB in India

In terms of absolute number, India has the most TB of any country. In 2020 there were an estimated 2.6 million cases of TB, 26% of the world's cases, with 494,000 TB-associated

deaths. A gap between incident cases and case notifications persist, although it had narrowed in recent years until 2020 (Figure 2.5) (26). Much of the burden is in children, and around one in eight total cases are extrapulmonary in nature, the result of *M. tuberculosis* disseminating from the lungs into another part of the body (23). There have been successful efforts to improve access to molecular testing in recent years (27), but rapid molecular testing is still only used to identify about 18% of TB cases (26). Standardised patient studies have found that providers infrequently suggest sputum-based TB testing, including molecular tests (28, 29).

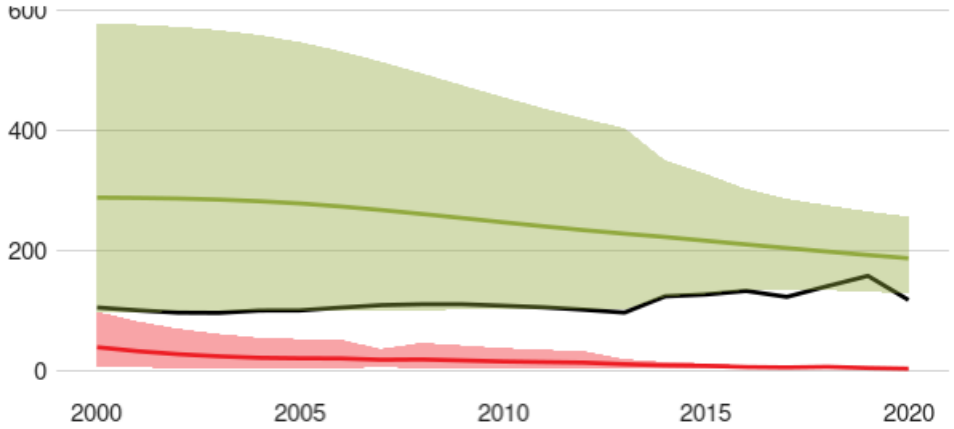


Figure 2.5: TB cases in India from 2000 to 2020. The green line indicates the number of incident TB cases, with shading indicating uncertainty. The black line indicates the total number of TB cases notified to the National TB Program. The red line represents incident TB cases in people living with HIV. All figures are per 100,000 (26).

2.1.4 TB in Peru

Peru has been designated by WHO as a high TB burden country, specifically because of it high per capital levels of multi-drug resistant TB (MDR-TB) (22). Although incident cases have been in decline for the past two decades, in 2020, there were still an estimated 38,000 TB cases (Figure 2.6), about 1700 of whom were started on treatment for drug-resistant TB (DR-TB) (30). Peru has implemented molecular testing networks in several cities across the country and private laboratories are also beginning to offer molecular testing. In Lima, the capital city, the GeneXpert platform is available for TB testing in 16 reference laboratories (31). However, WHO reports that only 5.2% of notified TB cases were initially diagnosed with a molecular test.

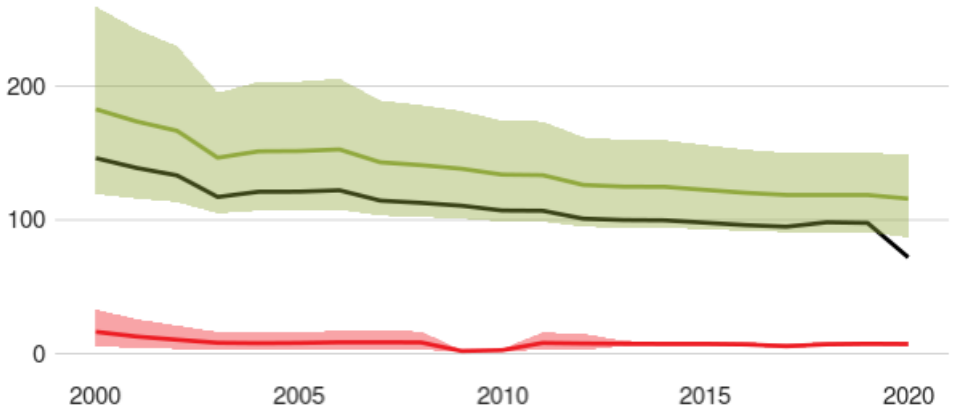


Figure 2.6: TB cases in Peru from 2000 to 2020. The green line indicates the number of incident TB cases, with shading indicating uncertainty. The black line indicates the total number of TB cases notified to the National TB Program. The red line represents incident TB cases in people living with HIV. All figures are per 100,000 (30).

2.1.5 Tuberculosis and COVID-19: an unfolding crisis

The COVID-19 pandemic has reversed hard-won gains in the fight against TB. One year into the pandemic, countries were reporting drops in TB case notifications ranging from 16% to 41% (mean 23%), observed incidence levels not reported since 2008 (1). This is due to a combination of lockdowns, reallocation of existing public health tools and personnel to the COVID-19 response, and patient fears of accessing healthcare. Some evidence suggests that the risk of acquiring new *M. tuberculosis* infections likely declined due to widespread uptake of masking and physical distancing; however, increased duration of close household contact may mean that the risk of transmitting infection in these settings is actually higher. Therefore, while certain pandemic-related public health interventions have likely been helpful in limiting the spread of TB, other accompanying phenomenon may erase these benefits (32). Factors leading to changes in the TB cascade of care are shown in Figure 2.7. As people with TB who develop COVID-19 are at an increased risk of severe disease and mortality compared to those without COVID-19, finding people with TB and ensuring they receive care is essential (33, 34).

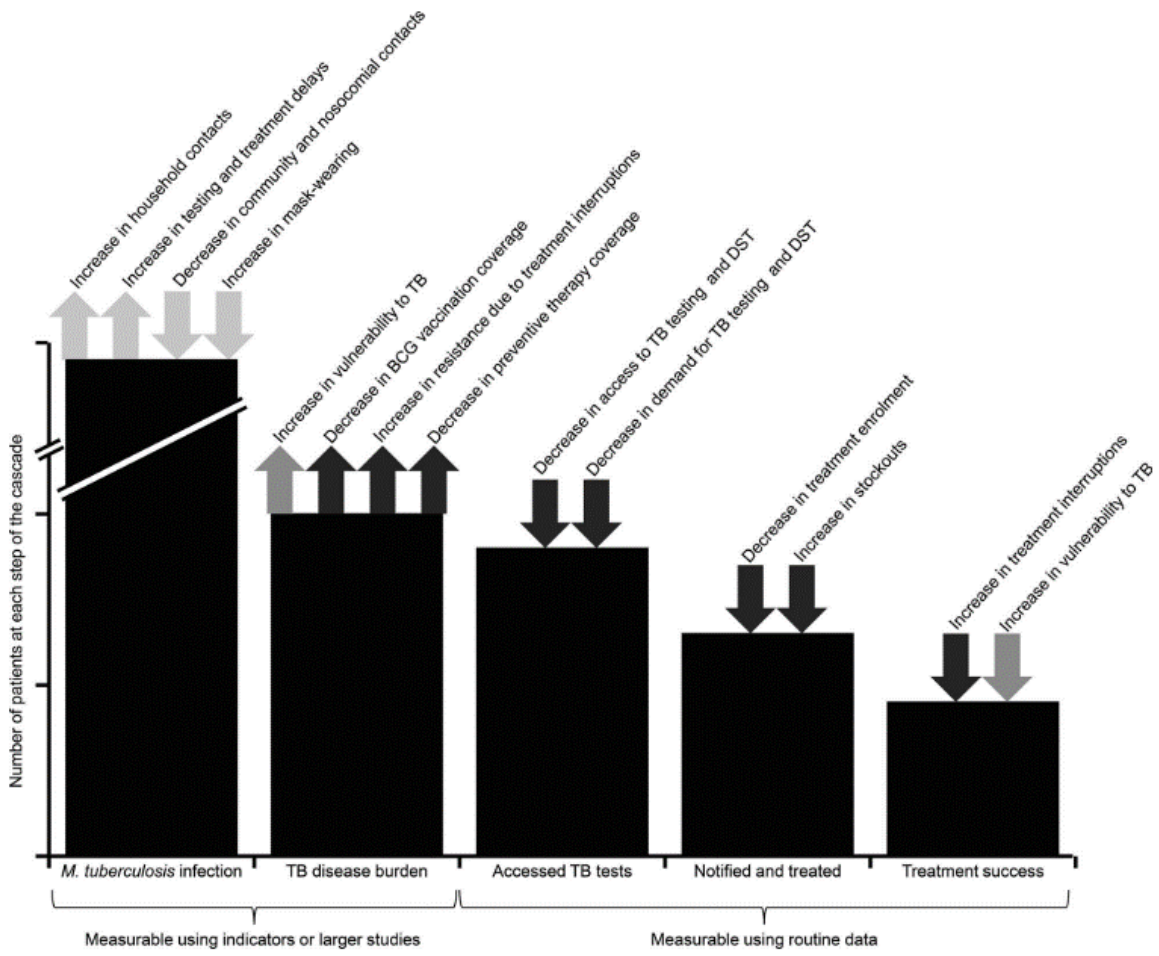


Figure 2.7: Potential impacts of COVID-19 on the TB care cascade. Arrows indicate an increase or decrease in number of patients at that step of the cascade. Black arrows indicate an impact on health service delivery and demand, grey arrows indicate an impact on vulnerability to TB, and light grey arrows indicate an impact on M. tuberculosis transmission (32).

In Peru, the COVID-19 pandemic has been devastating. Since 6 March 2020, when the first case was detected (35), there have been approximately 3.5 million cases and 212,000 deaths attributable to COVID-19 (36). Following trends observed elsewhere, TB case notifications dropped precipitously. From March to October 2020, 8093 fewer people were diagnosed than the corresponding period in 2019, a drop of 20% (Figure 2.8).

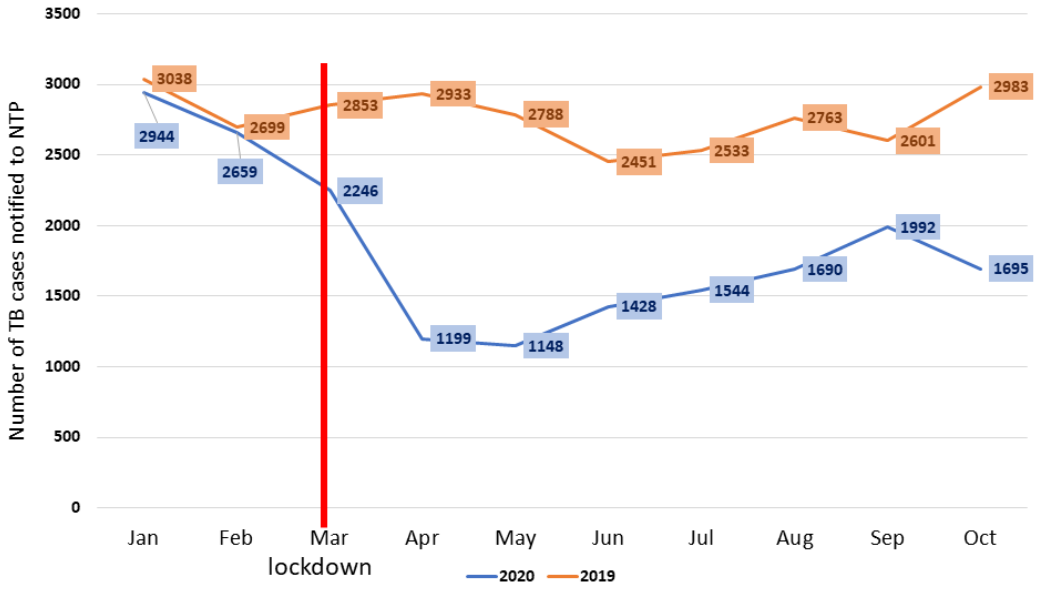


Figure 2.8: TB case notification changes in Peru, 2019-2020. The orange and blue figures are the numbers of TB cases notified to the Peruvian National TB Program in 2019 and 2020, respectively. The vertical red line indicates the time when COVID-19 restrictions were implemented. NTP – National TB Program (37).

However, permanent large losses in case notifications are not inevitable. In the state of Kerala, India, after a 50% drop in case notifications from April to July 2020, 90% of expected cases were recovered at year's end through intensified active case finding, symptom screening for TB in all people with COVID-19 followed by TB testing with Xpert, and use of a hub and spoke model for sample collection from every village (38). By prioritizing the continuation of molecular testing, the majority of TB cases in Kerala were found and connected to proper care. This is a template that other high-TB burden settings could emulate to recover lost TB cases. The rapid response to COVID-19 shows that action can be taken against infectious diseases like TB (39).

2.1.6 Integrated testing for TB and COVID-19

In an attempt to combat the decline in case notifications, some global TB stakeholders have begun to share testing policy recommendations for TB and COVID-19. Stop TB Partnership and USAID published a policy brief in early 2021 stating that, "[when] a person presents to a healthcare facility or provider with respiratory symptoms including cough and difficulty breathing, diagnostic tests for both COVID-19 and TB should be done at the same time (simultaneous testing) on a multiplex testing platform (integrated testing)" (40). They also note that individuals with physical risk factors for severe TB or COVID-19, such as diabetes or chronic respiratory disorders, should receive special attention for integrated testing (40). These recommendations are applicable to a broad population and capture anyone with presumptive TB or COVID-19.

Later in 2021, the Global Fund also published a briefing note providing guidance on testing for TB and COVID-19. Their recommendations apply to more specific populations. In particular,

they recommend that “[in] people who meet case definitions for both TB and COVID-19, systematic testing for both pathogens is recommended” when the diseases are prevalent in an individual’s community (41). They also recommend molecular testing in those who have had symptoms for longer than 7 days, and antigen rapid testing for those with symptoms lasting 5-7 days (Figure 2.9).

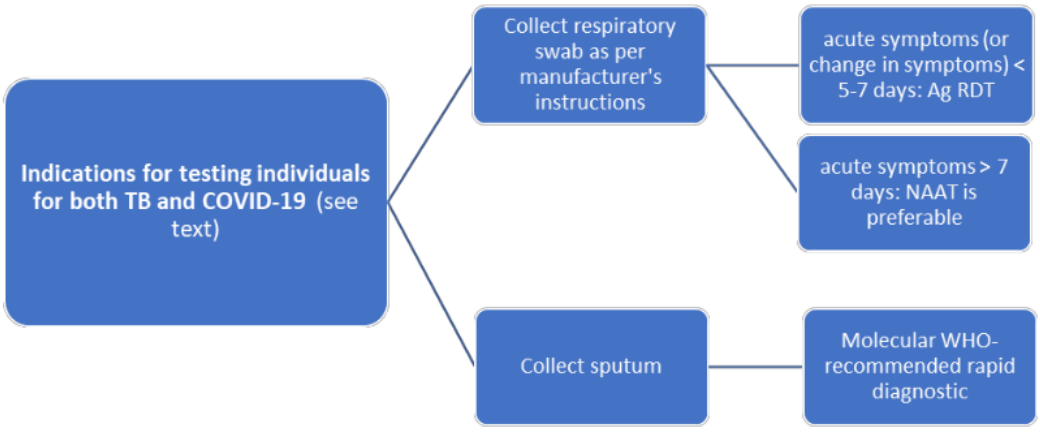


Figure 2.9: Global Fund's algorithm for testing for both TB and SARS-CoV-2 (41)

This testing algorithm is decidedly less broad than the catch-all approach proposed by Stop TB Partnership and USAID, but ultimately may not lead to such a difference in practice, as TB and COVID-19 share many symptoms. It can be difficult to discern between the two diseases based on clinical picture alone.

The Global Fund brief also suggests action for when people with TB are suspected to have developed COVID-19 (Figure 2.10). Should this situation arise, TB must continue to be properly managed, and then testing for SARS-CoV-2 infection may proceed if individuals “meet the COVID-19 case definition or when there is persistence or worsening of their condition despite appropriate treatment for the specific form of TB” (41).

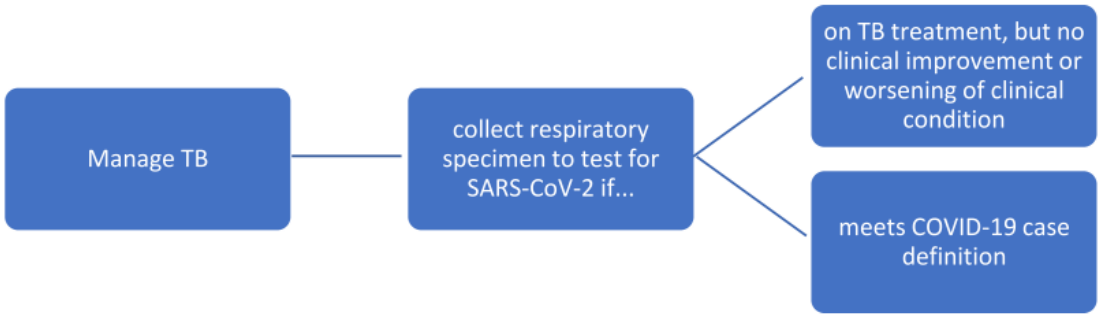


Figure 2.10: Indications for testing people with confirmed TB for COVID-19 (41)

There is great interest in integrated TB and COVID-19 testing as a solution to begin to recover the missing TB cases lost during the COVID-19 pandemic, and in the absence of WHO policy, other stakeholders have proposed guidance. However, knowledge gaps remain regarding this approach (42, 43). The feasibility of implementing integrated testing within existing laboratory and specimen collection workflows is unknown, as is the role of alternate specimens, such as sputum. The difference between the guidance issued by Stop TB Partnership/USAID and the Global Fund demonstrates that it is not yet well-understood who exactly should be prioritized for integrated testing: the target population needs to be more well-defined. The impact of integrated testing, with respect to its diagnostic yield, is unknown, and cost and cost-effectiveness have yet to be investigated. Ostensibly, integrated testing for TB and COVID-19 seems to be a beneficial intervention, but it has yet to be systematically evaluated and data are lacking.

2.2 Tuberculosis – diagnosis

2.2.1 New century, same tests

Since *M. tuberculosis* was identified by Robert Koch in 1882, TB has been diagnosed by microbiological methods, namely, sputum smear microscopy and mycobacterial culture. These assays continue to be commonly used today, even as a range of newer molecular tests have become available and received WHO-endorsement (8). Established and emerging molecular tests for TB are reviewed extensively in Chapter 3 (Manuscript 1). Diagnostic delays of weeks to months are not uncommon when undergoing testing for TB (44), and contribute to losses of patients from the TB care cascade. These delays can be partially attributed to the low accuracy of smear microscopy and the slow turn-around time of culture, the two tests that continue to function as the backbone of TB testing programs in many countries (9).

Smear microscopy's sensitivity ranges from 32-94%, depending on the patient's bacterial load and the skill and experience of the technician performing the test (45). Smear microscopy also cannot differentiate between drug-sensitive TB and DR-TB. Thus, although smear microscopy may quickly produce results, it often produces false negative results and drug-resistant cases may be started on inappropriate treatment. Culture is considered to be the gold standard method to diagnose TB and can detect drug-resistance, but even newer systems like BACTEC MGIT (BD, USA) require weeks to months to produce results (46), as well as trained laboratory staff and reference-level laboratories (47). Therefore, although culture is a critical tool in the TB diagnostic armamentarium, it is often too slow to immediately impact clinical decision-making.

The continued use of some microbiological testing is inevitable, particularly phenotypic drug susceptibility testing (DST) in the context of new types of DR-TB (48). However, for each microbiological test there is a faster, more sensitive molecular alternative. Molecular testing has been shown to reduce diagnostic delays and time to treatment initiation (49-52). Unfortunately, molecular assays continue to be prohibitively expensive in many high-TB burden settings and typically require significant investment in technology and training of personnel (27, 53-55).

Diagnostics developers need to work with country governments to arrive at solutions that will facilitate the increased deployment of molecular TB tests (56). Denying proper care to people with presumptive TB because they cannot afford a molecular test is not a defensible position.

Existing molecular tests have been optimized for use in detecting adult PTB with sputum samples. However, as mentioned above, substantial proportions of TB are found in children, and extrapulmonary disease forms are not rare. Children are often paucibacillary, meaning their disease is the result of low bacterial burden, making detection of *M. tuberculosis* challenging, and often they cannot readily produce sputum samples (57). For extrapulmonary TB, the area of the body suspected to be infected by *M. tuberculosis* must be sampled and tested; sampling of lymph nodes, ascitic fluid, cerebral spinal fluid, or elsewhere is invasive and requires a trained healthcare worker and resultant samples may be paucibacillary (58). Newer molecular tests with good accuracy that is well-characterised for all forms of TB in all populations could help close gaps in the TB care cascade.

2.3 Diagnostic accuracy

2.3.1 Diagnostic accuracy's connection to TB test scale-up

A critical step in the developmental pipeline for any new diagnostic test is the establishment of its diagnostic accuracy. Unfortunately, obtaining regulatory approval for diagnostic assays does not follow a procedure as formulaic as the approval process for treatments, i.e., a series of four hierarchical phases that gradually build evidence. In general, analytical validity and strong clinical validity, indicating a test's repeated ability to correctly classify individuals as diseased or non-diseased with high accuracy, are sufficient information to obtain regulatory approval (59, 60); however, this information does not need to be known for all possible test use cases (60). For TB diagnostics, this has resulted in tests that have undergone rigorous investigations into their ability to diagnose PTB (which has typically been presented as the highest interest use case), but the evidence of their accuracy in diagnosing TB in other forms or populations is often delayed or of lower certainty (8, 61). This can contribute to delays in scale-up and roll-out of tests for other non-PTB use cases, as countries often wait for WHO endorsement before updating their own policies (9, 62).

2.3.2 Characterising diagnostic accuracy

Conducting trials to characterise a diagnostic test's performance is often considered a hassle by test developers (59, 63), and studies producing accuracy estimates frequently suffer from limitations in design or analysis. These limitations ultimately lead to low confidence in findings, restricted generalizability, and uncertainty regarding applicability (59). This trend was recently observed in the beginning of the COVID-19 pandemic, when new tests to detect SARS-CoV-2 were often evaluated in poorly-designed studies (63, 64) or were compared to other new tests

with essentially unknown accuracy, with the resultant percent agreement termed “sensitivity” or “specificity” (65).

Sensitivity and specificity, however, are markedly different from percent agreement. Percent agreement is simply the percentage of tests run that produce concurring results. Positive percent agreement indicates the percentage of the tests that concurrently produce positive results, while negative percent agreement indicates the percentages of tests that concurrently produce negative results. In either case, there is no indication of whether the agreement is correct or incorrect. Contrastingly, to compute sensitivity or specificity, a comparison needs to be made between the new test of interest (the “index test”) and some sort of reference standard. The reference standard defines whether the person undergoing testing has the target condition, typically a disease, or not. Sensitivity, then, is the proportion of subjects with a positive result out of all people with the target condition; in other words, it is the proportion of individuals with the target condition who have a positive result on the index test. Specificity is the proportion of subjects with a negative result out of all people without the target condition, or the proportion of individuals without the target condition who have a negative result on the index test (66).

2.3.3 Composite reference standards

When a reference standard with perfect accuracy for a particular disease exists, estimating sensitivity and specificity is a straight-forward calculation. Unfortunately, many diseases are without a perfect reference standard test (67), including TB, which complicates the characterization of a test’s diagnostic accuracy. A comparison to an imperfect test would lead to biased estimates of the index test’s accuracy. In such cases, it is common to construct a composite reference standard (CRS) that combines the results from multiple imperfect tests using a particular decision rule. Other components such as clinical signs or epidemiological history may also be included. Disease status is defined by a rule, for example an “and” rule, where all tests much be positive, or an “or” rule, where any one test result much be positive. This is a satisfying approach as sensitivity and specificity can then be calculated based on study participants’ diagnosed disease status as defined by the CRS, and understanding the rules is straight-forward and intuitive (68, 69). For example, the diagnostic accuracy of a new molecular test for TB in people living with HIV could be estimated by comparing it to a CRS of Xpert and culture, where a positive result on either Xpert or culture indicates TB-positivity.

However, the same issue can arise as when utilizing a single imperfect reference test: assuming the CRS is 100% accurate when really it is not will lead to biased estimates of the index test’s accuracy (70). Interestingly, the addition of more tests to the CRS may, in fact, lead to more biased results than a comparison of index test to one imperfect reference test (71). And unlike the situation with a single imperfect test, a multi-test CRS may be subject to conditional dependence between component and index tests, which may further bias results. Between-test conditional dependence arises when test results are associated with a variable other than true disease status; this may be a shared mechanism or the tests may tend to make the same errors if, for example, both assays are more likely to detect disease in individuals with a higher

bacterial or viral load (70-72). This can lead to over- or underestimates of the index test's accuracy (72-74). Consider that as soon as one component test in a CRS has imperfect specificity, some truly disease-free individuals will be misclassified as disease-positive, which will result in an under-estimation of the index test's sensitivity.

2.3.4 Latent class analysis

Latent class analysis (LCA) can overcome many of the issues that arise when utilizing a CRS. LCA is a statistical modeling technique that classifies people into latent, or unobserved and unknown, categories (75). Relatively recently, it has been deployed to estimate diagnostic accuracy and disease prevalence in the absence of a perfect reference standard, with "disease status" being the latent category of interest (76-78).

The latent class model produces probabilities of belonging to a particular disease class for each individual, explicitly incorporating the uncertainty introduced by using imperfect tests. By utilizing the test results of all available tests, incorporating between test conditional dependence (79, 80), and, if adopting a Bayesian approach, specifying relevant prior information (81), LCA can produce less biased estimates of sensitivity and specificity.

In LCA, the joint results from multiple available tests are modelled in terms of each test's imperfect sensitivity and specificity and the disease prevalence (and conditional dependences if relevant). Adjusted sensitivities and specificities for all tests are thus simultaneously estimated. The simplest case of the joint distribution of k tests (T_k), where i = 1, …, N denotes the subjects, with binary results (T_k=1 and T_k=0) and covariate vector X_i, for a 2-class (d_i) model can be parameterized as the following:

$$P(T_{1i},\ \ldots,\ T_{ki}) = \sum_{d_i=0}^{1} P(d_i|X_i = x_i) \times P(T_{1i},\ \ldots,\ T_{ki}|d_i)\ .$$

Other paramaterisations are certainly possible. Conditional dependence may be included in different ways, for example by including a covariate (82) or addition of a random effect (83). A random effect model may be an appropriate approach when the covariance structure is not known, such as when the conditional dependence is due to an unmeasured factor (83), for example, unmeasured *M. tuberculosis* burden.

~

As described above in the Objectives, in this thesis I have employed different methods for estimating diagnostic accuracy depending on the specific population of interest. The following manuscripts report my attempts to utilize these methods to produce accurate performance estimates, and in doing so, contribute to the evidence base supporting the expanded use of molecular TB testing.

3 CHAPTER 3: MANUSCRIPT 1: ADVANCES IN MOLECULAR DIAGNOSIS OF TUBERCULOSIS

3.1 Preface

In 2008, WHO recommended a molecular TB test for the first time. Since then, a variety of medical diagnostics companies have entered the space, with a growing range of products becoming available. As the field has expanded, some tests have received WHO endorsement and policies surrounding molecular TB testing have been updated. Trends in the space are changing quickly. Synthesising the recent developments is a valuable exercise to bring health practitioners and researchers up to speed.

In this manuscript, I review the state of the field of molecular TB diagnostics. I report the performance characteristics of both WHO-endorsed and emerging molecular assays. New trends and developments in the space are described and put into context of the larger TB epidemic.

This work was published in September 2020 in *Journal of Clinical Microbiology*. Relevant developments since publication are described in the Afterward section.

3.2 Title page

Advances in Molecular Diagnosis of Tuberculosis

Emily MacLean[a,b], Mikashmi Kohli[a,b], Stefan F. Weber[c], Anita Suresh[d], Samuel G. Schumacher[d], Claudia M. Denkinger[b,c], Madhukar Pai[a,b,e]

[a] Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

[b] McGill International TB Centre, McGill University, Montreal, Canada

[c] Department of Infectious Diseases, University of Heidelberg, Heidelberg, Germany

[d] Foundation for Innovative New Diagnostics, Geneva, Switzerland

[e] Manipal McGill Program for Infectious Diseases, Manipal Academy of Higher Education, Manipal, India

3.3 Abstract

Molecular tests for tuberculosis (TB) have the potential to help reach the three million people with TB who are undiagnosed or not reported each year and to improve the quality of care TB patients receive by providing accurate, quick results, including rapid drug-susceptibility. The World Health Organization (WHO) has recommended the use of molecular nucleic acid amplification tests (NAATs) for TB detection instead of smear microscopy, as they are able to detect TB more accurately, particularly in patients with paucibacillary disease and in people living with HIV. Importantly, some of these WHO-endorsed tests can detect mycobacterial gene mutations associated with anti-TB drug resistance, allowing clinicians to tailor effective TB treatment. Currently, a wide array of molecular tests for TB detection is being developed and evaluated, and while some tests are intended for reference laboratory use, others are being aimed at the point-of-care and peripheral healthcare settings. Notably, there is an emergence of molecular tests designed, manufactured, and rolled out in countries with high TB burden, of which some are explicitly aimed for near-patient placement. These developments should increase access to molecular TB testing for larger patient populations. With respect to drug susceptibility testing, NAATs and next generation sequencing can provide results substantially faster than traditional phenotypic culture. Here, we review recent advances and developments in molecular tests for detecting TB as well as anti-TB drug resistance.

3.4 Introduction

With an estimated 1.5 million attributable deaths and 10 million new cases in 2018, tuberculosis (TB) is the leading infectious disease killer globally (1). Despite the severity of the epidemic, approximately 3 million people with TB were deemed “missing” due to underdiagnosis as well as underreporting to national TB programs (1). The World Health Organization (WHO) End TB Strategy calls for finding these missing millions in order to meet the Sustainable Development Goal of ending TB by 2030. New diagnostic tests and optimized test deployment strategies will be critical to achieving this target (2). In the context of the ongoing COVID-19 pandemic, it is also important to consider integrating testing for TB as well as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), since symptoms and testing technologies overlap (3).

Over the last decade, the field of TB diagnostics has seen advances in the form of new molecular tests. Often referred to as nucleic acid amplification tests (NAATs), these assays rely on amplification of a targeted genetic region of *Mycobacterium tuberculosis* complex, typically by PCR. NAATs can detect TB and perform drug susceptibility testing (DST) for key drugs, such as rifampicin (RIF) and isoniazid (INH), more quickly than conventional mycobacterial culture and are also available at different levels of healthcare systems. As such, they are disrupting the field of TB diagnostics and are helping to improve the quality of TB care (4, 5). Here, we review recent advances in the field of molecular diagnostics for TB and relevant WHO policies and describe the emerging landscape. For advances in biomarker-based tests for active and latent TB detection, we refer the readers to other review articles (6, 7).

3.4.1 State of the art

As shown in Figure 3.1 and Table 3.1, there are several molecular TB tests that are already WHO recommended and commercially available. Since Xpert MTB/RIF (Cepheid, Sunnyvale, USA) was first endorsed in 2010, advances in the field of TB diagnostics have mostly been in the realm of NAATs and responsive to the needs articulated by published target product profiles (TPPs) (8, 9). More than ever before, new assays are emerging and undergoing validation for TB and TB drug-resistance detection. However, simply developing new tests is insufficient for ensuring their implementation in countries with the highest TB burdens, and barriers to scale-up molecular tests like Xpert MTB/RIF have been identified (10). A 2018 study showed that despite high diagnostic accuracy and quick time to results, the ratio of smear microscopy tests to Xpert tests performed in 17 countries with a high TB burden was 6 to 1 (11). A similar trend of low uptake of new TB tests has also been reported for urine lipoarabinomannan (LAM) testing (12). For new tests to have impact, they must be adopted and scaled up (13).

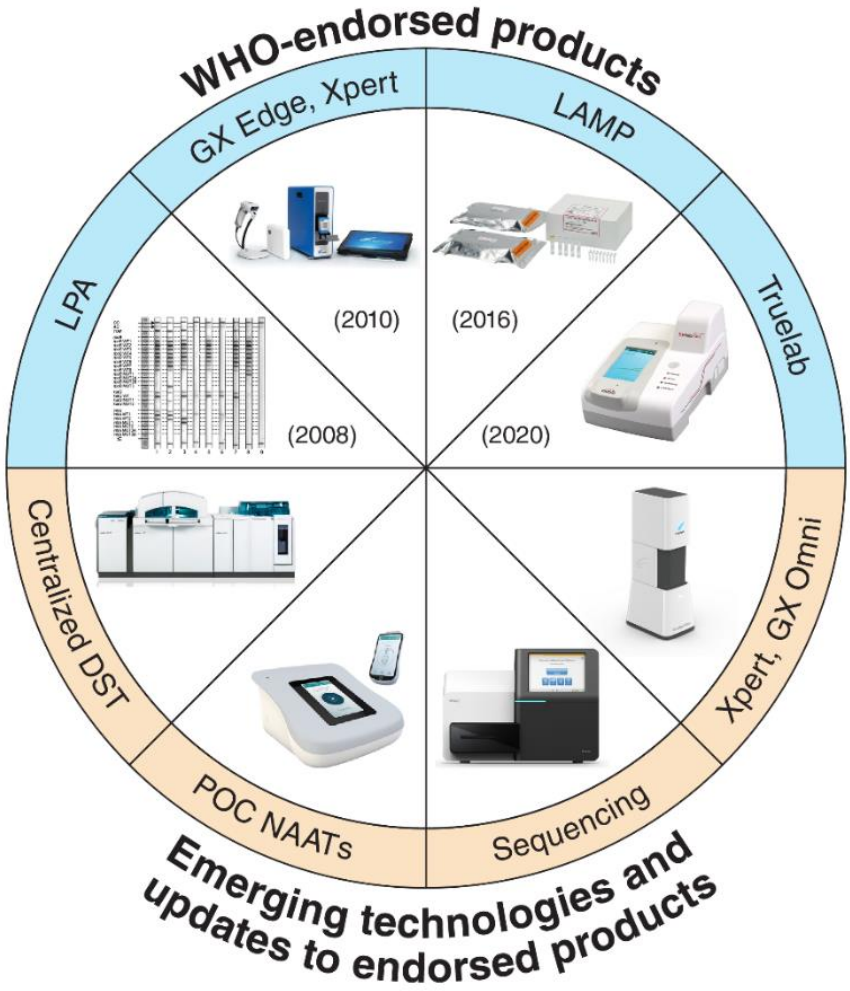


Figure 3.1: WHO-endorsed and emerging molecular tests for TB and drug resistance. WHO-endorsed and emerging molecular tests for TB and drug resistance. Outlined in blue are WHO-endorsed NAATs, including LPAs (14), Xpert Ultra (20), LAMP (16), and Truelab (21). Tests that are not yet WHO-endorsed, but under development or evaluation, are outlined in orange. Images shown are examples of products within each category. DST – drug sensitivity testing; GX – GeneXpert; LAMP - loop-mediated isothermal amplification; LPA – line probe assay; POC – point of care; NAAT – nucleic acid amplification test.

3.5 Developments in test and platforms with WHO endorsement

Table 3.1 provides an overview of all currently available NAATs that are endorsed by WHO, along with information on diagnostic accuracy.

Table 3.1: WHO-endorsed molecular tests for pulmonary TB detection and drug susceptibility testing

Technology	Xpert MTB/RIF	Xpert MTB/RIF Ultra	First-line Line probe assays (e.g., GenoType MTBDRplus and NIPRO)	Second-line Line probe assays (e.g., GenoType MTBDRsl)	Loopamp MTBC assay	TrueNAT MTB Plus	TrueNAT MTB – RIF Dx
Year endorsed	2010	2017	2008	2016	2016	2020	2020
Method principle	qPCR	qPCR / melting temperature analysis (RIF resistance)	PCR, hybridization	PCR, hybridization	loop-mediated isothermal amplification	micro RT-PCR	micro RT-PCR
Intended use	MTB diagnosis & RIF resistance detection	MTB diagnosis & RIF resistance detection	Diagnosis of RIF & INH resistance	Diagnosis of FLQ & SLID resistance	MTB diagnosis	MTB diagnosis	Diagnosis of RIF resistance
Sensitivity*	85% (pooled) 96% (RIF resistance)	90% (pooled) 94% (RIF resistance)	98% (RIF resistance) 84% (INH resistance)	86% (FLQ resistance) 87% (SLID resistance)	78% (pooled)	89% (pooled)	93% (RIF resistance)
Specificity*	99% (MTB detection) 98% (RIF resistance)	96% (MTB detection) 98% (RIF resistance)	99% (RIF resistance) >99% (INH resistance)	99% (FLQ resistance) 99% (SLID resistance)	98% (MTB detection)	98% (MTB detection)	95% (RIF resistance)
Target setting of use	District or sub-district laboratory	District or sub-district laboratory	Reference laboratory	Reference laboratory	Peripheral laboratory	Peripheral laboratory	Peripheral laboratory
Turn-around time	<2 hours	<2 hours	5 hours	5 hours	<2 hours	< 2 hours	<2 hours
Amenable to rapid test-and-treat?	Yes, especially on Omni platform	Yes, especially on Omni platform	No	No	Yes	Yes, on Truelab platform	Yes, on Truelab platform
Reference for policy guidance	WHO 2020 (21) WHO 2016 (85)	WHO 2020 (21)	WHO 2008 (14)	WHO 2016 (15)	WHO 2016 (16)	WHO 2020 (21)	WHO 2020 (21)

FLQ – fluoroquinolone. INH – isoniazid. LAMP – loop-mediated isothermal amplification. MTB – Mycobacterium tuberculosis. NAAT – nucleic acid amplification tests. RIF – rifampicin. SLID – second line injectable drugs. SSM+/C- – sputum smear microscopy positive / culture positive.

SSM-/C+ - sputum smear microscopy negative / culture positive. WHO – World Health Organization.

***n.b.: performance estimates in Table 3.1 have been retrieved from different studies, and are not the result of head to head comparisons. Therefore, comparing performances between tests must be made with caution. All reported values are from the policy guidance document cited.**

3.5.1 Line probe assays

Line probe assays (LPA) for first-line TB drugs (INH and RIF) have been endorsed by WHO for over a decade for the detection of multiple-drug-resistant TB (MDR-TB) (14). These assays include GenoType MTBDRplus (Hain Lifesciences-Bruker, Nehren, Germany) and Nipro NTM+MDRTB II (Osaka, Japan). New-generation LPAs have emerged with higher sensitivity, and some (e.g. GenoType MTBDRsl Version 2.0; Hain Lifesciences-Bruker) can detect mutations associated with fluoroquinolones (FLQs) and second-line injectables, kanamycin, amikacin, and capreomycin, and are recommended to guide MDR-TB treatment initiation (15).

3.5.2 Loop-mediated isothermal amplification

Loop-mediated isothermal amplification (LAMP) is an isothermal PCR amplification technique that can be performed in peripheral healthcare settings. The LAMP-based TB-LAMP assay (Eiken Chemical Company, Tokyo, Japan) has been recommended by WHO as a potential replacement for smear microscopy since 2016, owing to its superior diagnostic performance. It also does not require much sophisticated laboratory equipment (16) (Table 3.1). Despite this, TB-LAMP is underutilized (17), but some countries are creating their own LAMP assays for in-country use. Hopefully, country-specific versions of LAMP will increase uptake.

3.5.3 Next generation Xpert testing

In 2010, WHO endorsed Xpert MTB/RIF use with the GeneXpert platform (Cepheid, Sunnyvale, USA (18), and an updated policy was released in 2013 (19). In 2017, WHO recommended Xpert Ultra (Cepheid) (Ultra) the next generation of Xpert MTB/RIF, as the initial TB diagnostic test for adults and children, regardless of HIV status, over smear microscopy and culture (20). As in previous generations, Ultra detects RIF resistance by employing four probes with targets in the *rpoB* gene and melting temperature analysis (Table 3.1). Compared to previous generations, Ultra test cartridges have a larger chamber for DNA amplification than Xpert MTB/RIF and two multicopy amplification targets for TB, *IS6110* and *IS1081*, for a lower limit of detection of 16 CFU/ml. These modifications have increased Ultra's overall sensitivity from 85% (95% confidence interval [CI], 82% to 88%) to 88% (95% CI, 85% to 91%); however, compared to the

previous generation, Ultra's specificity is lower at 96% (95% CI, 90% to 98%) versus 98% (95% CI, 97% to 98%), seemingly because it detects nonviable bacteria, particularly in people with recent TB (21, 22). This lower specificity is proving to be an important issue in certain settings, such as areas with high numbers of HIV-TB coinfections or recurrent TB cases, like South Africa. In a recent study by Mishra and colleagues, it was shown that the Xpert Ultra assay had significantly lower specificity and positive predictive value than the Xpert MTB/RIF assay and high numbers of Ultra positive/culture negative people with previous treatment (23). The clinical consequences of treating such patients are unclear, and ongoing studies are attempting to shed light on this information.

The Xpert Ultra test also has a semi-quantitative "trace" category, indicating bacilli at the lowest limits of detection. In instances of trace positives (termed "trace calls"), one of the two multicopy amplification targets, but not the *rpoB* sequences, are detected. In instances of suspected extrapulmonary TB, children, and people living with HIV (PLHIV), trace positives should be treated as positives, as these cases tend to be paucibacillary. For other cases, a fresh specimen should be retested to rule out false positives (20). Trace calls may be difficult to interpret, as in the aforementioned study by Mishra et al, where it was observed that among people who were previously treated for TB, trace positives were a substantial portion of all positives, and these individuals by definition had indeterminate results for RIF resistance and were culture-negative, precluding further DST (23). Trace calls may be improving Ultra's sensitivity for extrapulmonary TB, particularly in the context of definite or probable TB meningitis, where sensitivity of 70% (95%CI, 47% to 87%) in cerebrospinal fluid was observed (24); however this finding is not consistent across studies, as another group observed a sensitivity of definite or probable TB of only 49% (95%CI, 35% to 63%) (25). Notably, even with sensitivity of 77%, as observed in another study of TB meningitis (26), the Ultra test's negative predictive value is still too low for use as a rule-out test. Research on Ultra for TB lymphadenitis (27) has shown sensitivities of 70% using fine needle aspiration and 67% using tissue biopsy in a study of 99 people with suspected TB lymphadenitis (27). In a multisite study using 317 frozen pleural fluid samples, Ultra sensitivity was 44%, compared to Xpert at 19% (28). More research will be necessary to determine if Ultra's performance for other forms of extrapulmonary TB has improved over the Xpert MTB/RIF assay (29).

As an automated PCR-based test, Ultra can be used by minimally trained technicians, but as it runs on the GeneXpert platform, it requires a continuous power supply and computer which

limits its use as a true point-of-care (POC) test. Alternatively, the recently launched GeneXpert Edge system is battery-powered and utilizes a tablet, making it more portable.

3.5.4 Made in India: Truelab by Molbio

Truenat MTB, Truenat MTB Plus, and Truenat MTB-Rif Dx (Molbio Diagnostics, Goa, India) are chip-based, micro real-time PCR-based assays for TB detection that produce results in 1 hour on the portable Truelab platform (Molbio Diagnostics). Already being rolled out in India, Truenat is characterized as a more affordable alternative to Xpert that is made in India. Products that are developed and manufactured in a country with a high TB burden might be quicker and more straightforward to scale-up in that country than products developed in another country, as governments often already have a degree of buy-in, data from locally-run studies will have accumulated, and supply chain and regulatory issues are simpler to solve (30, 31).

Truenat MTB and Truenat MTB Plus detect MTB bacilli in sputum after extraction using the separate TruePrep instrument and kits, with Truenat MTB-Rif Dx available as an optional add-on chip for sequential RIF resistance detection (32). Truelab, which comes in Uno, Duo, and Quattro throughput formats, was designed to be “rugged” and POC friendly: it has a dust filter and runs in temperatures up to 30°C, but multiple micro-pipetting steps necessitate a trained technician for its operation.

In December 2019, WHO convened a guideline development group meeting to determine recommended use cases for Truenat assays and other rapid molecular tests. The subsequent rapid communication reported that Truenat MTB, MTB Plus, and MTB-Rif Dx displayed comparable sensitivities and specificities to Xpert MTB/RIF and Ultra for detection of TB and RIF resistance, although this is based on an interim analysis of a multi-centre study that is still on-going. A formal update of molecular TB testing which will include Truenat tests is expected in this year. The 2020 WHO Consolidated Guidelines on Molecular Diagnostics recommend using Truenat MTB or MTB Plus rather than smear microscopy as an initial diagnostic test for TB in adults and children with signs and symptoms of pulmonary TB. This is a conditional recommendation, as test accuracy certainty is moderate. Regarding DST, with a Truenat MTB- or MTB Plus-positive result, Truenat MTB-RIF Dx may be used as an initial test for RIF resistance rather than phenotypic DST. This is also a conditional recommendation, as there is very low certainty of evidence for test accuracy (21).

3.6 Emerging technologies

3.6.1 Xpert XDR

Another PCR-based cartridge has been designed to run on the GeneXpert and Omni platforms for the simultaneous detection of mutations associated with resistance to multiple first- and second-line TB drugs, or extensively drug resistant TB (XDR-TB). Against phenotypic drug-susceptibility testing, a prototype version of the Xpert XDR cartridge displayed sensitivities (95%CI) of 83.3% (77.1% to 88.5%) for INH, 88.4% (80.2% to 94.1%) for ofloxacin, 96.2% (87.0% to 99.5%) for moxifloxacin at a critical concentration of 2.0 µg per milliliter, 71.4% (56.7% to 83.4%) for kanamycin, and 70.7% (54.5% to 83.9%) for amikacin (33). In July 2020, the Xpert MTB XDR-TB cartridge was launched, but further validation and WHO review are pending (34). As WHO updates treatment guidelines for MDR-TB and XDR-TB, it will be critical that molecular tools for DST can be updated to quickly reflect new recommendations. Already, this iteration of Xpert XDR may have less impact than it otherwise would have, as WHO has de-emphasized second-line injectable agents in treating drug resistant forms of TB (35). Future developments will need to focus on drugs that are now critical for MDR and XDR-TB management, including bedaquiline, pretomanid, and linezolid (36), but developing highly accurate molecular diagnostics to detect resistance to these drugs is currently impossible due to the lack of knowledge on resistance mechanisms.

3.6.2 GeneXpert Omni and other point-of-care devices

The GeneXpert platform was originally designed for use at the district or subdistrict level. Although efforts were made to use the technology at lower tiers of the health system, it soon became evident that microscopy centres in countries with a high TB burden often lacked the infrastructure necessary for this technology, including continuous power and temperature controls (10). As such, the POC GeneXpert Omni platform is a long-awaited development, as it will permit the use of Xpert MTB/RIF and Ultra assays in decentralized locations (e.g. primary care centers). Although delays have pushed back its launch repeatedly, Omni promises to be a real POC platform with a 2-day battery life and no tablet or computer requirement (37). The first instruments will be available in 2021, and Omni will eventually be able to run Ultra and any other Xpert cartridges that become available.

Other such POC NAATs are also under development. For example, Q-POC from QuantuMDx (Newcastle-upon-Tyne, United Kingdom) is a POC battery-operated PCR system that promises to deliver TB testing results in less than 30 minutes. It has been evaluated in combination with oral swabs as a sample, where its sensitivity and specificity, in preliminary studies, were similar to that of Xpert (38).

3.6.3 Indigenous Chinese diagnostics

Similar to Molbio in India, Chinese biotechnology firms have used their own expertise to develop TB NAATs for in-country use. These companies have undergone the China Food and Drug Administration (CFDA) regulatory processes, received approval, and rolled out the tests nationally. However, none of these technologies have been reviewed by WHO, and therefore, uptake by other countries is limited. Table 3.2 summarizes the performance of some of these assays from systematic reviews (39, 40).

Table 3.2: CFDA-endorsed molecular test for TB diagnosis and drug susceptibility testing

Technology	EasyNAT	SAT-TB	MeltPro TB	GeneChip MDR
Method principle	Cross priming amplification	Isothermal amplification of MTB 16S RNA	PCR, melt curve analysis	PCR, hybridization
Intended use	MTB diagnosis	MTB diagnosis	DST	MDR-TB diagnosis; INH, RIF resistance
Sensitivity	87% (pooled)	71% - 94% [range]	98% (RIF resistance) 85% (INH resistance) 64% (FLQ resistance) 83% (SLID resistance)	79% (MDR-TB) 89% (RIF resistance) 79% (INH resistance)
Specificity	97% (pooled)	54% – 83% [range]	97% (RIF resistance) 98% (INH resistance) 98% (FLQ resistance) 99% (SLID resistance)	98% (MDR-TB) 97% (RIF resistance) 97% (INH resistance)
Target setting of use	District or sub-district laboratory	District or reference laboratory	Reference laboratory	Reference laboratory
Reference	Deng 2019 (39)	Deng 2019 (39)	Sun 2019 (40)	Sun 2019 (40)

CFDA – China Food and Drug Administration. DST – drug susceptibility testing. INH – isoniazid. RIF – rifampicin. SLID – second line infectable drugs.

CFDA-approved since 2014, EasyNAT (Ustar Biotechnologies, Hangzhou, China) replicates and detects mycobacterial DNA from sputum via cross-priming amplification (CPA). As CPA is an isothermal technique, Easynat may be placed at low levels of health care systems, as a thermal cycler is not required (39). A fully integrated and automated next generation version is in development (41).

Simultaneous amplification and testing (SAT)-TB (Rendu Biotechnology, Shanghai, China) detects Mycobacterial 16S rRNA from sputum, which is isothermally amplified before the resultant cDNA is detected by fluorescent probes, requiring laboratory infrastructure, such as adequate biosafety facilities for specimen manipulation and trained personnel (42).

For drug resistance testing, MeltPro TB (Zeesan Biotech, Xiamen, China) assays for RIF, INH, second-line injectables, and FLQs are available, allowing them to detect MDR-TB and XDR-TB. After manual DNA extraction, MeltPro TB detects drug resistance via melt curve analysis using a PCR machine; the shift in melting temperature from wild type to mutation in sequences covered by multiple probes can be qualitatively detected (43).

GeneChip MDR (CapitalBio Corporation, Beijing, China) is a microarray assay that requires hands-on sample preparation before reverse hybridization and analysis on a fully automated system. As such, it requires sophisticated laboratory equipment. GeneChip MDR utilizes multiplexed asymmetric PCR to detect resistance to RIF and INH in one assay, and thus can detect MDR-TB (44).

3.6.4 High-throughput solutions: centralized diagnostic tests

Recently, centralized, high-throughput NAATs for TB diagnosis and drug resistance detection have been developed and are currently undergoing WHO evidence evaluation. RealTime MTB (Abbott Molecular, Abbott Park, USA), RealTime RIF/INH (Abbott Molecular), FluoroType MTB (Hain Lifescience, Nehren, Germany), FluoroType MTDBR (Hain Lifescience), Cobas MTB (Roche, Rotkreuz, Switzerland), and Max MDR-TB (BD, Franklin Lakes, USA) assays run on established multi-disease platforms that are already employed for such diseases as human immunodeficiency virus (HIV), human papillomavirus, and hepatitis C virus (45). These almost entirely automated tests are all intended for tertiary laboratory use. In 2019, a WHO technical expert group meeting reported that the centralized assays' performance for detecting resistance to INH and RIF was similar to LPA and that RealTime MTB, Cobas MTB, and Max MDR-TB performed similarly to Xpert MTB/RIF for TB detection (46). For now, these assays are recommended for operational research use only, with a WHO review of broader use expected in late 2020.

The RealTime MTB is a multiplex NAAT that targets the MTB *IS6100* and *PAB* genes with a limit of detection (LOD) of 17 CFU/ml. Up to 96 respiratory specimens can be inactivated and processed by the Abbott m2000 platform per run (47). A systematic review and meta-analysis of ten studies incorporating 4858 specimens found that RealTime MTB had a sensitivity of 96%

(95% CI, 90% to 99%) and specificity of 97% (95% CI, 94% to 99%) for TB detection; regarding RIF resistance detection, it had a pooled sensitivity of 94% (95% CI, 89% to 99%) and specificity of 100% (95% CI, 99% to 100%); and for INH resistance detection, its pooled sensitivity was 89% (95% CI, 86% to 92%) and specificity was 99% (95% CI, 98% to 100%) (45).

Another centralized test is the semi-automated FluoroType MTB, a beacon-based PCR assay performed on the Hain Fluorocycler platform. Specimen decontamination, sample preparation, and DNA isolation must be performed manually, which requires 30 minutes of hands-on time, with the entire process taking 4 h to final results (48). In a systematic review and meta-analysis of five studies incorporating 2660 specimens, FluoroType MTB displayed a sensitivity of 92% (95% CI, 88% to 93%) and specificity of 99% (95% CI, 64% to 100%) (45).

The Cobas 6800/8800 MTB assay runs on the high throughput Cobas 8800 platform that can run 960 samples in 8 hours. One internal manufacturer study of 744 samples reported a sensitivity and specificity of 95% (95% CI, 92% to 97%) and 98% (95% CI, 96% to 99%), respectively (49).

Finally, the Max MDR-TB test runs on the BD Max platform and targets the MTB 16S rRNA gene. Up to 24 specimens are manually decontaminated and prepared before extraction and amplification by the Max MDR-TB assay. Time to final results is 4 h (50). A manufacturer-sponsored validation study of 892 samples reported TB detection sensitivity of 93% (95% CI, 89% to 96%) and specificity of 97% (95% CI, 96% to 98%). Sensitivity for RIF resistance and INH resistance were 90% (95% CI, 55% to 100%) and 82% (95%CI, 63% to 92%), respectively, with 100% specificity in both cases (51).

Centralized TB assays are promising due to their high diagnostic accuracy and ability to run large numbers of samples simultaneously, and their automated nature reduces the hazard of contacting infectious respiratory specimens for healthcare workers and laboratory technicians. The developmental pipeline for centralized assays is quite robust, with platforms, such as MeltPro (Zeesan Biotech), Seegene (Seoul, South Korea), and MolecuTech (YD Diagnostics, Seoul, South Korea) currently under regulatory assessment (52). All platforms are offering tests for MDR-TB and XDR-TB, which will provide further options in the future.

However, carry-over contamination is still possible with these assays, and quality assurance is critical. Additionally, the costs for each of these tests have not been made public, and no subsidized or concessional pricing schemes are yet in place. These tests do run on multi-

disease platforms, which adds value, but it is unclear exactly who will be willing to pay to implement these tests if they can only perform DST for INH and RIF resistance, particularly when there are simpler NAATs available (Table 3.1). Furthermore, their centralized placement means they are unavailable where patients first present to care, and therefore, sample transportation is essential for success. Reliable systems for delivering test results to patients and healthcare providers must also be in place for these tests to have impact.

3.6.5 Next-generation sequencing

Next-generation sequencing (NGS) is increasingly considered a promising option for comprehensive DST for TB and produces results much faster than traditional phenotypic culture or culture-based testing (53, 54). Unlike probe-based assays where detection is limited to probe-specific targets, NGS-based assays can provide detailed and accurate sequence information for whole genomes, as with whole genome sequencing (WGS), or multiple gene regions of interest, as with targeted NGS (55) (Table 3.3).

Table 3.3: Strengths and limitations of WGS versus targeted sequencing via next generation sequencing

Whole genome sequencing	Targeted sequencing
Strengths • Full genome sequenced • No pre-specified targets needed • Comprehensive solution • Detect rare mutations and hetero-resistance	**Strengths** • Sequence directly from sample • Large number of gene targets • Less expensive than WGS • Simpler bioinformatics and storage • Detect rare mutations and hetero-resistance
Weaknesses • Requires culture isolates • Slower than targeted NGS • Complicated bioinformatics • Expensive	**Weaknesses** • Knowledge of targets required • Less information than WGS • Expensive

Acknowledging the value of NGS, WHO has published guidance on the role of sequencing for detecting mutations associated with drug resistance in TB (55), along with a consensus-based TPP for sequencing. In 2019, a TB sequencing database called ReSeqTB was established at

WHO to curate, standardize, and unify genotypic and phenotypic DST data, along with metadata on DR-TB (56).

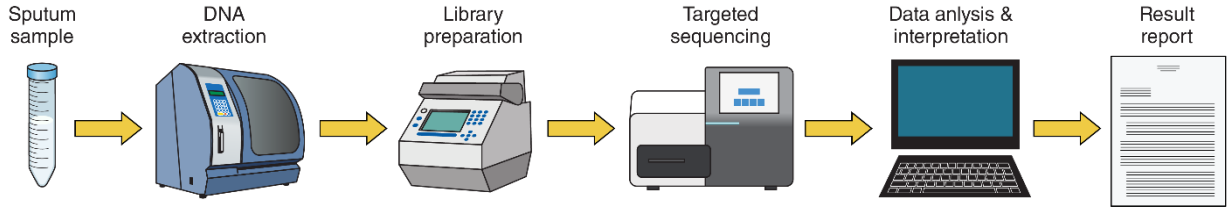


Figure 3.2: Targeted sequencing workflow schematic

There are ongoing efforts by multiple stakeholders to validate targeted sequencing as a complete end-to-end solution for DR-TB detection, from DNA extraction direct from respiratory samples (i.e., without the need for first culturing and then isolating a specimen), targeted library preparation and sequencing, to result reporting (Figure 3.2). One such targeted assay that is currently available in the market is Deeplex Myc-TB (Genoscreen, Lille, France). Deeplex Myc-TB uses ultra-deep sequencing of 24-plex amplicon mixes for mycobacterial species identification, genotyping, and DST. In addition, the manufacturer indicates that it can detect hetero-resistance, i.e., the phenomenon of subpopulations within a seemingly uniform microbial population displaying both resistance and susceptibility to a particular drug (57), down to 3% of minority strains in cases of multiple infections or emergent mutations (58). Another newly-developed targeted sequencing assay for DR-TB is DeepChek-TB (Translational Genomics Research Institute, Flagstaff, USA), which has recently been licensed by ABL (Luxembourg) (59). Both tests are currently for research use only.

Sequencing is currently being successfully implemented for DR-TB surveillance purposes in at least seven countries - Azerbaijan, Bangladesh, Belarus, Pakistan, Philippines, South Africa, and Ukraine (60). Select health systems in settings with low TB burden, including the United Kingdom (Public Health England), the Netherlands, and New York state, have already transitioned from phenotypic culture to WGS for DST for first-line drugs (61, 62). The US Centers for Disease Control and Prevention sequence isolates from all culture-confirmed TB cases nationwide (63).

More countries are considering switching to a sequencing-based approach for surveillance of drug susceptibility. For example, India has recently expressed interest in utilizing sequencing for surveillance and clinical care. In 2018, infrastructure and technical support for sequencing was introduced at five National TB Program laboratories across India with Global Fund funding. It is

hoped that this will be the beginning of the foundations of a clinical diagnostic network in the future (64).

South Africa has implemented and integrated sequencing into their national drug resistance surveillance program as an alternative to phenotypic DST and are considering its future potential for laboratory-based TB management and TB transmission investigations (65).

In Brazil, the interdisciplinary group Rede Brasileira de Pesquisas em Tuberculose (REDE-TB, Brazilian TB Research Network) identified NGS as a key technology for implementation. Through the Oswaldo Cruz Foundation (Fiocruz), Brazil has also signed memoranda of understanding with the Beijing Genomic Institute and the Chinese Centre for Disease Control and Prevention. One of the planned activities under this agreement is the establishment of a sequencing service at Fiocruz with applications in infectious disease, including TB (66).

Regarding sequencing for DST, centralized sequencing platforms have been the norm, but there is increasing interest in smaller and more portable sequencing devices, such as MinION (Oxford Nanopore, Oxford, UK) (67) and iSeq from Illumina (San Diego, USA) (68), validation for both of which is on-going.

3.7 Potential for integrating NAAT testing for TB and COVID-19

Across the world, healthcare systems are being upended by the COVID-19 pandemic, but it is critical we do not neglect other diseases like TB that persist outside of the spotlight (69, 70). A modelling study suggests that one unintended result of the pandemic-related lockdowns is a projected 1.5 million excess TB-related deaths from 2020-2025 (3). Countries must act now to ensure routine care for patients experiencing other disease continues and to ensure these projections to not become reality.

One clear area for intervention is the integration of TB and COVID-19 testing. As patients with either disease may present with cough, fever, or difficulty breathing, this represents an opportunity to test presumptive patients for TB and COVID-19 in one clinical encounter. This would be more convenient for patients and healthcare workers as it could reduce the number of necessary follow-up visits.

The recently launched Xpert Xpress SARS-CoV-2 (Cepheid) cartridge might allow low- and middle-income countries to increase their capacity to test for COVID-19, as many countries already have existing GeneXpert networks (71). However, concern has been expressed that a ramp-up of COVID-19 testing on the GeneXpert system may come at the expense of TB testing in low- and middle-income countries (LMICs) that rely on Xpert MTB/RIF (3, 69). Abbott and

Roche also have released COVID-19 assays to run on their centralized testing platforms, RealTime (72) and Cobas 6800 (73), respectively. Both systems are used in some reference laboratories of countries with high TB burden for multi-disease testing. In India, Molbio has released a COVID-19 test for the Truelab platform (74) that is now in use.

Leveraging existing multi-disease NAAT platforms for both TB and COVID-19 testing could be an effective strategy. Research into using a common respiratory sample (e.g. sputum) will be necessary to understand feasibility of this strategy and to address biosafety concerns.

3.8 Conclusion: optimizing the impact of NAATs

Advances in molecular TB diagnostics in the last decade have resulted in TB tests that are highly accurate and faster than conventional microbiological tests, and emerging technologies promise to continue this trend. In some respects, NAATs are having positive clinical impact. For example, it has been shown that routine use of Xpert leads to reductions in time to TB diagnosis and time to treatment initiation, from several days to same-day (4, 5, 75-77), and its use also facilitated increased numbers of patients to commence anti-TB treatment (5, 77). However, for longer-term outcomes like mortality, NAAT impact is more ambiguous (78, 79) albeit inherently difficult to measure appropriately (80).

As long as cascades of care in high TB settings remain weak or fragmented, diagnostic testing alone will be unable to decrease mortality or disease recurrence, and evaluations of NAAT clinical significance will continue to produce null results (81). Issues such as under-utilization of existing NAATs, empirical treatment of people with suspected TB, and patient loss to follow-up all reduce the potential beneficial effect of diagnostic testing (82). Thus, optimizing the clinical impact of molecular tests for TB will require their introduction into functioning, strengthened healthcare systems, which can also respond to outbreaks that require multidisease testing capacity (83). Centering patients within high quality health systems will allow NAATs to reach their full potential and become an integral part of a digitally connected, patient-centered, re-imagined TB care system (84).

3.9 Afterward

Since this review was published, WHO has released updated consolidated guidelines for TB. Regarding initial testing for TB, the guidelines strongly recommend the use of Xpert for TB and RIF resistance detection for all adults and children with signs and symptoms of pulmonary TB. Xpert Ultra is strongly recommended as the initial diagnostic test in adults who have no prior history of TB, but is only presented as an option for those with recent (<5 years) history of TB or TB treatment due to the difficulty in interpreting trace calls. Xpert and Xpert Ultra are also strongly recommended for the initial diagnosis of presumptive TB meningitis in adults and children. For other forms of extrapulmonary TB, due to limited confidence in the evidence, Xpert and Ultra are presented as options for initial testing. Trunat MTB and MTB Plus continue to be conditionally recommended as initial diagnostic tests for TB diagnosis and rifampicin resistance detection in adult and children with signs and symptoms of pulmonary TB, but this recommendation does not extend to PLHIV (84). Unfortunately, the GeneXpert Omni system has been cancelled by Cepheid (85), leaving the need for a POC NAAT option unfulfilled.

There have been some notable updates in the drug resistance detection space. An updated TPP was issued for next-generation molecular DST in peripheral settings, laying out the specific characteristics the TB community is seeking over the next five years, with an emphasis on sequencing (86). It is hoped that the recent publication of a comprehensive M. tuberculosis mutation catalogue (87) will aid diagnostic developers in creating products needed to ensure universal DST becomes reality. In terms of specific assays, Xpert XDR has been launched, but the definition of extensively drug resistant TB (XDR-TB) has since been updated (88), so it remains to be seen how much impact this assay will have. Recent data suggests that Xpert XDR is very specific for all drugs in its purview and its sensitivity was similar to that of line probe assays, although Xpert XDR sensitivity was higher for INH resistance (89).

Additionally, the centralized, high-throughput NAATs (termed “moderate complexity automated NAATs” by WHO) described in the review have now been recommended for the diagnosis of pulmonary TB in adults, including PLHIV, with TB signs and symptoms. They have also been endorsed for the detection of RIF and INH resistance detection. These recommendations are being extended to children based on generalisation of the existing adult data (84).

Building on momentum in the molecular diagnostics space due to COVID-19, more new molecular TB tests are in the developmental pipeline. A growing number of test manufacturers have platforms that are now well-established thanks to the pandemic, and work is underway to develop TB tests for these systems (85). Specific examples include the battery-powered Lumira Dx instrument, which has seen wide uptake across Africa since the COVID-19 pandemic’s onset and can produce test results in 20 minutes (90) and SD Biosensor’s point-of-care STANDARD M10 system, which can run runs both PCR- and LAMP-based assays (91). TB detection assays for these multiplex, peripheral setting-friendly platforms are forthcoming. Hopefully, the now widely-understood importance of diagnostics translates into sustained funding and political will to increase scale-up of molecular TB tests.

3.10 References

1. World Health Organization. 2019. Global tuberculosis report 2019. World Health Organization, WHO Press, Geneva.
2. World Health Organization. 2015. WHO End TB Strategy. World Health Organization, Geneva.
3. Stop TB Partnership. 2020. The potential impact of the COVID-19 response on tuberculosis in high-burden countries: a modelling analysis. Stop TB Partnership, Geneva.
4. Boyd R, Ford N, Padgen P, Cox H. 2017. Time to treatment for rifampicin-resistant tuberculosis: systematic review and meta-analysis. Int J Tuberc Lung Dis 21:1173-1180.
5. Theron G, Zijenah L, Chanda D, Clowes P, Rachow A, Lesosky M, Bara W, Mungofa S, Pai M, Hoelscher M, Dowdy D, Pym A, Mwaba P, Mason P, Peter J, Dheda K. 2014. 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. Lancet 383:424-35.
6. MacLean E, Broger T, Yerlikaya S, Fernandez-Carballo BL, Pai M, Denkinger CM. 2019. A systematic review of biomarkers to detect active tuberculosis. Nat Microbiol 4:748-758.
7. Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, Metcalfe JZ, Cattamanchi A, Dowdy DW, Dheda K, Banaei N. 2014. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. Clin Microbiol Rev 27:3-20.
8. Denkinger CM, Kik SV, Cirillo DM, Casenghi M, Shinnick T, Weyer K, Gilpin C, Boehme CC, Schito M, Kimerling M, Pai M. 2015. Defining the Needs for Next Generation Assays for Tuberculosis. Journal of Infectious Diseases 211:S29-S38.
9. Denkinger CM, Dolinger D, Schito M, Wells W, Cobelens F, Pai M, Zignol M, Cirillo DM, Alland D, Casenghi M, Gallarda J, Boehme CC, Perkins MD. 2015. Target Product Profile of a Molecular Drug-Susceptibility Test for Use in Microscopy Centers. Journal of Infectious Diseases 211:S39-S49.
10. Albert H, Nathavitharana RR, Isaacs C, Pai M, Denkinger CM, Boehme CC. 2016. Development, roll-out and impact of Xpert MTB/RIF for tuberculosis: what lessons have we learnt and how can we do better? Eur Respir J 48:516-25.
11. Cazabon D, Pande T, Kik S, Van Gemert W, Sohn H, Denkinger C, Qin ZZ, Waning B, Pai M. 2018. Market penetration of Xpert MTB/RIF in high tuberculosis burden countries: A trend analysis from 2014 - 2016. Gates Open Research In press:1-8.
12. Singhroy DN, MacLean E, Kohli M, Lessem E, Branigan D, England K, Suleiman K, Drain PK, Ruhwald M, Schumacher SG, Denkinger CM, Waning B, Van Gemert W, Pai M. 2020. Adoption and uptake of the lateral flow urine LAM test in countries with high tuberculosis and HIV/AIDS burden: current landscape and barriers [version 1; peer review: 2 approved]. Gates Open Res 4.
13. Pai M, Furin J. 2017. Tuberculosis innovations mean little if they cannot save lives. Elife 6.
14. World Health Organization. 2008. Molecular Line Probe Assays for Rapid Screening of Patients at Risk of Multi-Drug Resistant Tuberculosis: Policy Statement. World Health Organization, Geneva.
15. World Health Organization. 2016. The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs. World Health Organization, Geneva.
16. World Health Organization. 2016. The use of loop-mediated isothermal amplification (TB-LAMP) for the diagnosis of pulmonary tuberculosis: Policy guidance. World Health Organisation, Geneva.

17. Shete PB, Farr K, Strnad L, Gray CM, Cattamanchi A. 2019. Diagnostic accuracy of TB-LAMP for pulmonary tuberculosis: a systematic review and meta-analysis. BMC infectious diseases 19:268-268.
18. World Health Organization. 2010. WHO endorses new rapid tuberculosis test. World Health Organization, London | Geneva.
19. World Health Organization. 2013. Policy Update. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. World Health Organization, Geneva, Switzerland.
20. World Health Organization. 2017. Next-generation Xpert® MTB/RIF Ultra assay recommended by WHO. World Health Organization, Geneva.
21. World Health Organization. 2020. Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update. World Health Organization, Geneva.
22. Horne DJ, Kohli M, Zifodya JS, Schiller I, Dendukuri N, Tollefson D, Schumacher SG, Ochodo EA, Pai M, Steingart KR. 2019. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev 6:Cd009593.
23. Mishra H, Reeve BWP, Palmer Z, Caldwell J, Dolby T, Naidoo CC, Jackson JG, Schumacher SG, Denkinger CM, Diacon AH, van Helden PD, Marx FM, Warren RM, Theron G. 2020. Xpert MTB/RIF Ultra and Xpert MTB/RIF for diagnosis of tuberculosis in an HIV-endemic setting with a high burden of previous tuberculosis: a two-cohort diagnostic accuracy study. Lancet Respir Med doi:10.1016/s2213-2600(19)30370-4.
24. Bahr NC, Nuwagira E, Evans EE, Cresswell FV, Bystrom PV, Byamukama A, Bridge SC, Bangdiwala AS, Meya DB, Denkinger CM, Muzoora C, Boulware DR, Williams DA, Taseera K, Nyehangane D, Ivan M, Orikiriza P, Rhein J, Hullsiek KH, Musubire A, Pastick K, Nabeta P, Mwesigye J, Rajasingham R. 2018. Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: a prospective cohort study. The Lancet Infectious Diseases 18:68-75.
25. Donovan J, Thu DDA, Phu NH, Dung VTM, Quang TP, Nghia HDT, Oanh PKN, Nhu TB, Chau NVV, Ha VTN, Hang VTT, Trinh DHK, Geskus RB, Tan LV, Thuong NTT, Thwaites GE. 2020. Xpert MTB/RIF Ultra versus Xpert MTB/RIF for the diagnosis of tuberculous meningitis: a prospective, randomised, diagnostic accuracy study. The Lancet Infectious Diseases 20:299-307.
26. Cresswell FV, Tugume L, Bahr NC, Kwizera R, Bangdiwala AS, Musubire AK, Rutakingirwa M, Kagimu E, Nuwagira E, Mpoza E, Rhein J, Williams DA, Muzoora C, Grint D, Elliott AM, Meya DB, Boulware DR. 2020. Xpert MTB/RIF Ultra for the diagnosis of HIV-associated tuberculous meningitis: a prospective validation study. Lancet Infect Dis doi:10.1016/s1473-3099(19)30550-x.
27. Antel K, Oosthuizen J, Malherbe F, Louw VJ, Nicol MP, Maartens G, Verburgh E. 2020. Diagnostic accuracy of the Xpert MTB/Rif Ultra for tuberculosis adenitis. BMC infectious diseases 20:33-33.
28. Wang G, Wang S, Yang X, Sun Q, Jiang G, Huang M, Huo F, Ma Y, Chen X, Huang H. 2020. Accuracy of Xpert MTB/RIF Ultra for the Diagnosis of Pleural TB in a Multicenter Cohort Study. Chest 157:268-275.
29. Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, Steingart KR. 2018. Xpert((R)) MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. Cochrane Database Syst Rev 8:Cd012768.
30. Pai M. 17 February 2020 2020. Global health technologies: time to re-think the 'trickle down' model. Forbes.

31. Indian Council of Medical Research. 2017. India moving forward in research towards new tools for tuberculosis, p 1-2. Government of India, New Delhi.
32. Nikam C, Jagannath M, Narayanan MM, Ramanabhiraman V, Kazi M, Shetty A, Rodrigues C. 2013. Rapid diagnosis of Mycobacterium tuberculosis with Truenat MTB: a near-care approach. PLoS One 8:e51121.
33. Xie YL, Chakravorty S, Armstrong DT, Hall SL, Via LE, Song T, Yuan X, Mo X, Zhu H, Xu P, Gao Q, Lee M, Lee J, Smith LE, Chen RY, Joh JS, Cho Y, Liu X, Ruan X, Liang L, Dharan N, Cho SN, Barry CE, 3rd, Ellner JJ, Dorman SE, Alland D. 2017. Evaluation of a Rapid Molecular Drug-Susceptibility Test for Tuberculosis. N Engl J Med 377:1043-1054.
34. FIND. 2020. New rapid molecular test for tuberculosis can simultaneously detect resistance to first- and second-line drugs. FIND, Geneva.
35. World health Organization. 2019. WHO consolidated guidelines on drug-resistant tuberculosis treatment. World Health Organization, Geneva.
36. Conradie F, Diacon AH, Ngubane N, Howell P, Everitt D, Crook AM, Mendel CM, Egizi E, Moreira J, Timm J, McHugh TD, Wills GH, Bateson A, Hunt R, Van Niekerk C, Li M, Olugbosi M, Spigelman M. 2020. Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. New England Journal of Medicine 382:893-902.
37. Treatment Action Group. 2019. 2019 Pipeline Report: Tuberculosis Diagnostics. Treatment Action Group, Group TA, New York.
38. Ortega C, Wood R, Murton H, Andama A, Cattamanchi A, Dixon R, Morgan G, Madan D, Somoskovi A, Cangelose G. 2019. Diagnosis of pulmonary tuberculosis by oral swab analysis (OSA): optimisation and development of non-sputum, point-of-care methods. Int J Tuberc Lung Dis 23:S211.
39. Deng S, Sun Y, Xia H, Liu Z, Gao L, Yang J, Zhao Y, Huang F, Feng J, Wang L, Huan S, Zhan S. 2019. Accuracy of Commercial Molecular Diagnostics for the Detection of Pulmonary Tuberculosis in China: A Systematic Review. Sci Rep 9:4553.
40. Sun Y, Gao L, Xia H, Yang Z, Deng S, Yang J, Zhao Y, Wang L, Feng J, Huang F, Huan S, Zhan S. 2019. Accuracy of molecular diagnostic tests for drug-resistant tuberculosis detection in China: a systematic review. Int J Tuberc Lung Dis 23:931-942.
41. Unitaid. 2017. UNITAID end-of-project evaluation: TB GeneXpert – Scaling up access to contemporary diagnostics for TB. Unitaid, Geneva.
42. Yan L, Tang S, Yang Y, Shi X, Ge Y, Sun W, Liu Y, Hao X, Gui X, Yin H, He Y, Zhang Q. 2016. A Large Cohort Study on the Clinical Value of Simultaneous Amplification and Testing for the Diagnosis of Pulmonary Tuberculosis. Medicine (Baltimore) 95:e2597.
43. Pang Y, Dong H, Tan Y, Deng Y, Cai X, Jing H, Xia H, Li Q, Ou X, Su B, Li X, Zhang Z, Li J, Zhang J, Huan S, Zhao Y. 2016. Rapid diagnosis of MDR and XDR tuberculosis with the MeltPro TB assay in China. Scientific reports 6:25330-25330.
44. Guo Y, Zhou Y, Wang C, Zhu L, Wang S, Li Q, Jiang G, Zhao B, Huang H, Yu H, Xing W, Mitchelson K, Cheng J, Zhao Y. 2009. Rapid, accurate determination of multidrug resistance in M. tuberculosis isolates and sputum using a biochip system. Int J Tuberc Lung Dis 13:914-20.
45. Kohli M, MacLean E, Schumacher SG, Pai M, Denkinger CM. Under review. Diagnostic accuracy of centralized assays for TB detection and detection of resistance to rifampicin and isoniazid: A systematic review and meta-analysis. Eur Respir J.
46. World Health Organization. 2019. Evaluation of centralized assays for TB detection and detection of resistance to rifampicin and isoniazid: WHO Technical Expert Consultation Report. World Health Organization, Organization WH, Geneva.
47. Tang N, Frank A, Pahalawatta V, Lampinen J, Coblenz-Korte A, Dunn C, Li C, Cloherty G, Abravaya K, Leckier G. 2015. Analytical and clinical performance of Abbott RealTime

MTB, an assay for detection of Mycobacterium tuberculosis in pulmonary specimens. Tuberculosis (Edinb) 95:613-9.
48. Hofmann-Thiel S, Hoffman H. 2014. Evaluation of Fluorotype MTB for detection of Mycobacterium tuberculosis complex DNA in clinical specimens from a low-incidence country. BMC Infect Dis 14:1-5.
49. Roche Molecular Systems I. 2018. Cobas(R) MTB: Nucleic acid test for use on the Cobas(R) 6800/8800 systems. Roche Molecular Systems, Basel.
50. Rocchetti TT, Silbert S, Gostnell A, Kubasek C, Widen R. 2016. Validation of a Multiplex Real-Time PCR Assay for Detection of Mycobacterium spp., Mycobacterium tuberculosis Complex, and Mycobacterium avium Complex Directly from Clinical Samples by Use of the BD Max Open System. J Clin Microbiol 54:1644-7.
51. Shah M, Paradis S, Betz J, Beylis N, Bharadwaj R, Caceres T, Gotuzzo E, Joloba M, Mave V, Nakiyingi L, Nicol MP, Pradhan N, King B, Armstrong D, Knecht D, Maus CE, Cooper CK, Dorman SE, Manabe YC. 2019. Multicenter Study of the Accuracy of the BD MAX MDR-TB Assay for Detection of Mycobacterium tuberculosis Complex and Mutations Associated with Resistance to Rifampin and Isoniazid. Clin Infect Dis doi:10.1093/cid/ciz932.
52. FIND. 2020. Diagnostics Pipeline Tracker: Tuberculosis, *on* FIND. https://www.finddx.org/dx-pipeline-status/. Accessed
53. Lee RS, Pai M. 2017. Real-Time Sequencing of Mycobacterium tuberculosis: Are We There Yet? J Clin Microbiol 55:1249-1254.
54. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, Iqbal Z, Feuerriegel S, Niehaus KE, Wilson DJ, Clifton DA, Kapatai G, Ip CLC, Bowden R, Drobniewski FA, Allix-Béguec C, Gaudin C, Parkhill J, Diel R, Supply P, Crook DW, Smith EG, Walker AS, Ismail N, Niemann S, Peto TEA. 2015. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. The Lancet Infectious Diseases 15:1193-1202.
55. World Health Organization. 2018. Technical guide on next-generation sequencing technologies for the detection of mutations associated with drug resistance in Mycobacterium tuberculosis complex. World Health Organization, Organization WH, Geneva.
56. World Health Organization. 2019. Relational sequencing TB data platform [website]. World Health Organization, Geneva.
57. El-Halfawy OM, Valvano MA. 2015. Antimicrobial Heteroresistance: an Emerging Field in Need of Clarity. Clinical Microbiology Reviews 28:191-207.
58. Genoscreen. 2019. Deeplex(R) Myc-TB: From clinical samples to drug resistance profile. Genoscreen, Lille, France.
59. Yozwiak S. 2019. TGen and ABL plan global rollout of advanced test for TB, one of the world's most deadly pathogens. TGen, Flaggstaff, Arizona and Luxembourg City, Luxembourg.
60. Zignol M, Cabibbe AM, Dean AS, Glaziou P, Alikhanova N, Ama C, Andres S, Barbova A, Borbe-Reyes A, Chin DP, Cirillo DM, Colvin C, Dadu A, Dreyer A, Driesen M, Gilpin C, Hasan R, Hasan Z, Hoffner S, Hussain A, Ismail N, Kamal SMM, Khanzada FM, Kimerling M, Kohl TA, Mansjo M, Miotto P, Mukadi YD, Mvusi L, Niemann S, Omar SV, Rigouts L, Schito M, Sela I, Seyfaddinova M, Skenders G, Skrahina A, Tahseen S, Wells WA, Zhurilo A, Weyer K, Floyd K, Raviglione MC. 2018. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study. Lancet Infect Dis 18:675-683.
61. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, Farhat MR, Guthrie JL, Laukens K, Miotto P, Ofori-Anyinam B, Dreyer V, Supply P, Suresh A, Utpatel C, van Soolingen D, Zhou Y, Ashton PM, Brites D, Cabibbe AM, de Jong BC, de

Vos M, Menardo F, Gagneux S, Gao Q, Heupink TH, Liu Q, Loiseau C, Rigouts L, Rodwell TC, Tagliani E, Walker TM, Warren RM, Zhao Y, Zignol M, Schito M, Gardy J, Cirillo DM, Niemann S, Comas I, Van Rie A. 2019. Whole genome sequencing of Mycobacterium tuberculosis: current standards and open issues. Nat Rev Microbiol 17:533-545.
62. Shea J, Halse TA, Lapierre P, Shudt M, Kohlerschmidt D, Van Roey P, Limberger R, Taylor J, Escuyer V, Musser KA. 2017. Comprehensive Whole-Genome Sequencing and Reporting of Drug Resistance Profiles on Clinical Cases of Mycobacterium tuberculosis in New York State. J Clin Microbiol 55:1871-1882.
63. Armstrong GL, MacCannell DR, Taylor J, Carleton HA, Neuhaus EB, Bradbury RS, Posey JE, Gwinn M. 2019. Pathogen Genomics in Public Health. N Engl J Med 381:2569-2580.
64. Revised National Tuberculosis Control Programme. 2017. National Strategic Plan for Tuberculosis Elimination 2017–2025. Directorate General of Health Services, New Delhi.
65. National Institute for Communicable Diseases. 2018. National Institute for Communicable Diseases Annual Overview 2018/2019. National Institute for Communicable Diseases, Capetown.
66. Agência Fiocruz de Notícias. 2018. Fiocruz establishes partnership with Chinese scientific institutions. Oswaldo Cruz Foundation, São Paulo.
67. Cabibbe AM, Battaglia S, Spitaleri A, Suresh A, Colman RE, Uplekar S, Rodwell TC, Cirillo DM. 2019. A portable real-time solution for next generation sequencing-based diagnosis of drug-resistant tuberculosis direct from clinical samples. Int J Tuberc Lung Dis 23:S268.
68. Colman RE, Mace A, Seifert M, Hetzel J, Mshaiel H, Suresh A, Lemmer D, Engelthaler DM, Catanzaro DG, Young AG, Denkinger CM, Rodwell TC. 2019. Whole-genome and targeted sequencing of drug-resistant Mycobacterium tuberculosis on the iSeq100 and MiSeq: A performance, ease-of-use, and cost evaluation. PLoS Med 16:e1002794.
69. Venkatesan P. 2020. COVID-19 diagnostics—not at the expense of other diseases. The Lancet Microbe 1:e64.
70. Amimo F, Lambert B, Magit A. 2020. What does the COVID-19 pandemic mean for HIV, tuberculosis, and malaria control? Trop Med Health 48:32.
71. United States Food and Drug Administration. 2020. Coronavirus (COVID-19) Update: FDA Issues first Emergency Use Authorization for Point of Care Diagnostic. United States Food and Drug Administration, Atlanta.
72. Moore NM, Li H, Schejbal D, Lindsley J, Hayden MK. 2020. Comparison of two commercial molecular tests and a laboratory-developed modification of the CDC 2019-nCoV RT-PCR assay for the detection of SARS-CoV-2. J Clin Microbiol doi:10.1128/jcm.00938-20.
73. Broder K, Babiker A, Myers C, White T, Jones H, Cardella J, Burd EM, Hill CE, Kraft CS. 2020. Test Agreement Between Roche Cobas 6800 and Cepheid GeneXpert Xpress SARS-CoV-2 Assays at High Cycle Threshold Ranges. J Clin Microbiol doi:10.1128/jcm.01187-20.
74. Molbio Diagnostics Private Limited. 2020. New Launch: Truenat(TM) SARS CoV-2 Chip-based Real Time PCR Test for COVID-19. Molbio Diagnostics Private Limited,, Goa.
75. Lessells RJ, Cooke GS, McGrath N, Nicol MP, Newell ML, Godfrey-Faussett P. 2017. Impact of Point-of-Care Xpert MTB/RIF on Tuberculosis Treatment Initiation. A Cluster-randomized Trial. Am J Respir Crit Care Med 196:901-910.
76. Pereira GR, Barbosa MS, Dias NJD, Almeida CPB, Silva DR. 2018. Impact of introduction of Xpert MTB/RIF test on tuberculosis (TB) diagnosis in a city with high TB incidence in Brazil. PLoS One 13:e0193988.

77. Cox HS, Mbhele S, Mohess N, Whitelaw A, Muller O, Zemanay W, Little F, Azevedo V, Simpson J, Boehme CC, Nicol MP. 2014. 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 11:e1001760.
78. Auld AF, Fielding KL, Gupta-Wright A, Lawn SD. 2016. Xpert MTB/RIF - why the lack of morbidity and mortality impact in intervention trials? Trans R Soc Trop Med Hyg 110:432-44.
79. Di Tanna GL, Khaki AR, Theron G, McCarthy K, Cox H, Mupfumi L, Trajman A, Zijenah LS, Mason P, Bandason T, Durovni B, Bara W, Hoelscher M, Clowes P, Mangu C, Chanda D, Pym A, Mwaba P, Cobelens F, Nicol MP, Dheda K, Churchyard G, Fielding K, Metcalfe JZ. 2019. Effect of Xpert MTB/RIF on clinical outcomes in routine care settings: individual patient data meta-analysis. The Lancet Global health 7:e191-e199.
80. Schumacher SG, Denkinger CM. 2019. The impact of Xpert MTB/RIF-do we have a final answer? Lancet Glob Health 7:e161-e162.
81. Pai M, Schumacher SG, Abimbola S. 2018. Surrogate endpoints in global health research: still searching for killer apps and silver bullets? BMJ Glob Health 3:e000755.
82. Ochodo E, Kalema N, Schumacher S, Steingart K, Young T, Mallett S, Deeks J, Cobelens F, Bossuyt P, Nicol M, Cattamanchi A. 2019. Variation in the observed effect of Xpert MTB/RIF testing for tuberculosis on mortality: A systematic review and analysis of trial design considerations [version 1; peer review: 1 approved with reservations]. Wellcome Open Research 4.
83. Schumacher SG, Sohn H, Qin ZZ, Gore G, Davis JL, Denkinger CM, Pai M. 2016. Impact of Molecular Diagnostics for Tuberculosis on Patient-Important Outcomes: A Systematic Review of Study Methodologies. PLoS One 11:e0151073.
84. Stop TB Partnership, McGill International TB Centre, Arcady Group. 2020. Re-Imagining TB Care, *on* Stop TB Partnership. Accessed 5 April 2020.
85. World Health Organization. 2016. The use of the Xpert MTB/RIF assay for the diagnosis of TB. Meeting Report. World Health Organization, Geneva, Switzerland.

4 CHAPTER 4: MANUSCRIPT 2: DIAGNOSTIC ACCURACY OF STOOL XPERT MTB/RIF FOR THE DETECTION OF PULMONARY TUBERCULOSIS IN CHILDREN: A SYSTEMATIC REVIEW AND META-ANALYSIS

4.1 Preface

Children may have difficulty spontaneously producing sputum, the typical specimen used to diagnose TB, and the specimens that can be obtained frequently have low numbers of *M. tuberculosis*. This can make diagnosis of pediatric TB a challenge. Molecular diagnostic tools like Xpert MTB/RIF (Xpert) have a lower limit of detection than smear microscopy, but a sensitive test is of little use without an adequate sample. Therefore, interest has been growing in utilizing molecular tests with alternative, off-label specimens to detect TB in children.

In this manuscript, I systematically reviewed and meta-analysed the evidence on the use of Xpert with stool samples for the diagnosis of childhood TB for the first time. I estimated stool Xpert's sensitivity and specificity compared to microbiological reference standards and, when possible, clinical reference standards. I discuss the potential role of this test for diagnosing TB in children going forward and the short-comings in study quality and reporting in this area.

This work was published in May 2019 in *Journal of Clinical Microbiology*. Relevant developments since publication are described in the Afterward section.

4.2 Title page

Diagnostic accuracy of stool Xpert MTB/RIF for the detection of pulmonary tuberculosis in children: A systematic review and meta-analysis

Emily MacLean[a,b], Giorgia Sulis[a,b], Claudia M. Denkinger[b,c,d], James C. Johnston[b,e,f], Madhukar Pai[a,b], Faiz Ahmad Khan[a,b,g]

[a] Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Canada

[b] McGill International TB Centre, McGill University, Montreal, Canada

[c] Foundation for Innovative New Diagnostics, Geneva, Switzerland

[d] Department of Infectious Diseases, University of Heidelberg, Heidelberg, Germany

[e] Division of Respiratory Medicine, University of British Columbia, Vancouver, Canada

[f] BC Centre for Disease Control, Vancouver, Canada

[g] Respiratory Epidemiology & Clinical Research Unit, Research Institute of the McGill University Health Centre & Montreal Chest Institute, Montreal, Canada

4.3 Abstract

Invasive collection methods are often required to obtain samples for the microbiological evaluation of children with presumptive pulmonary tuberculosis (PTB). Nucleic-acid amplification testing of easier-to-collect stool samples could be a non-invasive method of diagnosing PTB. We conducted a systematic review and meta-analysis to evaluate the diagnostic accuracy of testing stool with the Xpert MTB/RIF assay ("stool Xpert") for childhood PTB. Four databases were searched for publications from January 2008 to June 2018. Studies assessing the diagnostic accuracy among children of stool Xpert compared to a microbiological reference standard of conventional specimens tested by mycobacterial culture or Xpert were included. Bivariate random-effects meta-analyses were performed to calculate pooled sensitivity and specificity of stool Xpert against the reference standard. From 1,589 citations, 9 studies (n=1,681) were included. Median participant ages ranged from 1.3 to 10.6 years. Protocols for stool processing and testing varied substantially, with differences in reagents and methods of homogenization and filtering. Against the microbiological reference standard, the pooled sensitivity and specificity of stool Xpert were 67% (95% confidence interval [CI], 52 to 79%) and 99% (95% CI, 98 to 99%), respectively. Sensitivity was higher among children with HIV (79% [95% CI, 68 to 87%] versus 60% [95% CI, 44 to 74%] among HIV-uninfected children). Heterogeneity was high. Data were insufficient for subgroup analyses amongst children under age 5, the most relevant target population. Stool Xpert could be a non-invasive method of ruling in PTB in children, particularly those with HIV. However, studies focused on children under 5 years of age are needed, and generalizability of the evidence is limited by the lack of standardized stool preparation and testing protocols.

4.4 Introduction

At least 1 million incident tuberculosis (TB) cases and 230,000 TB-related deaths are estimated to have occurred among children in 2017, accounting for approximately 10% of total cases and 15% of deaths (1). Pulmonary TB (PTB) is the most common form of childhood TB (2). Xpert MTB/RIF (Xpert) (Cepheid, USA), an automated cartridge-based PCR assay, is currently recommended by the World Health Organization (WHO) as the initial diagnostic test in presumptive PTB cases for adults and children (3). Minimal sample preparation is required, and test results are produced within 2 h. In a meta-analysis that pooled data from sputum smear-positive and -negative subjects, the performance of Xpert on respiratory samples had a sensitivity of 62% (95% credible interval, 51 to 73%) and specificity of 98% (95% credible interval, 97 to 99%). The use of Xpert on sputum is thus more sensitive than smear microscopy. Moreover, Xpert has several operational advantages over mycobacterial culture, the gold standard for TB diagnosis (4). However, in children under 5 years old, and particularly in those under 2 years old, the collection of sputum specimens for diagnostic testing is difficult and often requires invasive methods that are challenging to implement in resource-limited settings (e.g., nasopharyngeal/nasogastric aspiration or bronchoscopy) and not widely available (2). Furthermore, as pediatric TB is typically paucibacillary, the sensitivity of currently deployed tests is diminished in children versus adults (5).

Mycobacterium tuberculosis-containing sputum may be swallowed, particularly during sleep, and acid-fast bacilli have been shown to survive digestion and are detectable in stool (6, 7). As such, stool may represent a more acceptable and feasible alternative to conventional specimens for the evaluation of suspected childhood PTB. The use of Xpert on stool has not been included in recommendations by WHO, nor has any claim been made by the manufacturer regarding stool. However, several groups have now developed preprocessing methods in order to use Xpert on stool for the diagnosis of childhood TB.

We performed a systematic review and meta-analysis of the diagnostic performance of Xpert using stool samples for PTB in children.

4.5 Methods

4.5.1 Protocol and registration

The protocol for this systematic review and meta-analysis was registered at the International Prospective Register of Systematic Reviews (PROSPERO) (identifier CRD42017079836).

4.5.2 Search strategy and information sources

PubMed, EMBASE, Scopus, and the Cochrane Library were systematically searched from 1 January 2008 until 15 June 2018. The search strategy was developed with a medical librarian and based on key validated terms for “children” and “Xpert”, as well as “tuberculosis” with no filters applied. The full search strategies for each database are presented in Text S1 in the supplemental material. Experts in TB diagnostics were consulted to identify relevant papers that may have been missed by the search strategy. Citations of reviews and included publications were also searched.

4.5.3 Eligibility criteria

Publications in English, French, Italian, Mandarin, Spanish, and Portuguese; of any design and sampling strategy; and of any enrolment timing (prospective, retrospective, cross-sectional) were eligible for inclusion. Conference proceedings and abstracts, commentaries, editorials, and reviews were excluded, as were studies with a sample size less than 10. To be included, eligible studies must have reported the diagnostic performance of stool Xpert in patients under 16 years old, compared to a microbiological reference standard for the diagnosis of PTB. Studies that did not explicitly state that their focus was PTB were eligible if the types of specimens used for the reference standard were those that are typically used for PTB diagnosis (e.g. gastric aspirate). Studies that used banked sputum and stool specimens originally collected from children were also eligible.

4.5.4 Study screening and selection

Search results were imported into a citation manager and duplicates were removed. Two authors (E. MacLean and G. Sulis) independently screened citations by title and abstract per pre-defined eligibility criteria, followed by full-text review for all selected studies. Results disagreed upon were discussed, and a third reviewer consulted if necessary (F. Ahmad Khan).

4.5.5 Data extraction

A data extraction form was piloted by two reviewers (E. MacLean and G. Sulis) with critical input from a third (C. M. Denkinger). Two reviewers (E. MacLean and G. Sulis) independently extracted results using a standardized form (Text S2). After data extraction, results were compared, and disagreements were discussed until a consensus was reached. Study authors were contacted for missing performance data, clarification regarding reference standard definitions, and sample preparation techniques. Using these data and figures indicated in the

publications, we reconstructed two-by-two tables for stool Xpert performance compared to the microbiological reference standard and, where applicable, the clinical reference standard.

4.5.6 Risk-of-bias assessment

The Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) tool (8) was used to assess each included study's risk of bias. No formal assessment of publication bias was made, as traditional methods such as funnel plots and regression tests are not helpful for diagnostic studies (9).

4.5.7 Reference standards

Acceptable microbiological reference standards were mycobacterial culture or Xpert MTB/RIF, performed on specimens that are conventionally used to diagnose childhood PTB (nasogastric aspirates, gastric lavage, nasopharyngeal aspirates, and expectorated sputum). No studies included stool mycobacterial culture in their diagnostic work-up. Stool Xpert was not included in the reference standard.

Childhood PTB is often clinically diagnosed (i.e., without microbiologic confirmation). As such, we also examined the performance of stool Xpert compared to clinical reference standards that are compatible with updated international guidelines (5). Studies that followed these guidelines used a combination of signs and symptoms, chest radiography, epidemiological history, and tuberculin skin test (TST) results to classify children as "likely TB", "unconfirmed TB", and "unlikely TB" (Table S1). For our purposes, we dichotomized these outcomes into "likely/possible TB" and "unlikely TB".

4.5.8 Statistical Analysis

Data from reconstructed two-by-two tables were used to calculate sensitivity and specificity and associated 95% confidence intervals (CIs). In cases of empty cells in two-by-two tables, a zero correction was made by replacing the cell with 0.5. Aggregate-data meta-analyses were performed with bivariate random-effect hierarchical models (10) to estimate pooled sensitivity and specificity for stool Xpert compared to the microbiologic reference standard, and, separately, compared to the clinical reference standard. We also estimated pooled sensitivity and specificity stratified by HIV status. Results from individual studies and pooled estimates were presented on forest plots. To assess between-study heterogeneity, we used the I^2-statistic (11). In a sensitivity analysis, we estimated pooled sensitivity and specificity after excluding studies that used Xpert MTB/RIF but not mycobacterial culture of conventional specimens as

the microbiological reference standard. All analyses were conducted using the Midas package in STATA (STATA 15, Stata Corp., USA (12)). The study is reported following PRISMA guidelines (Table S2) (13).

4.6 Results

4.6.1 Search results

Our search identified 1,589 unique citations from which 34 studies were selected for full-text review, and 9 studies met inclusion criteria (Figure 4.1).

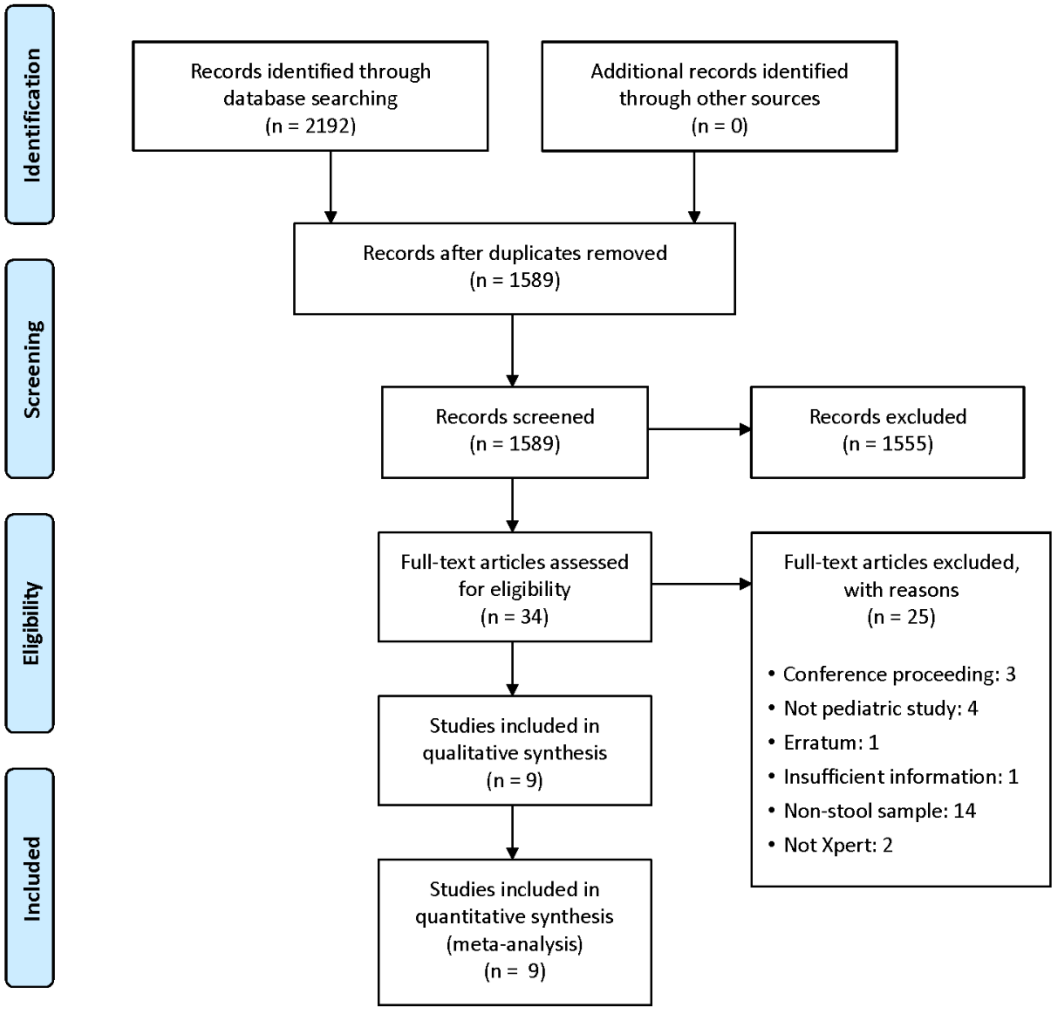


Figure 4.1: PRISMA study flow diagram

4.6.2 Study and participant characteristics

Study and patient characteristics are presented in Table 4.1. Among the 9 studies that we included, African countries were most well represented (7/9), whereas 2 studies recruited

participants from Asia. One study had multiple sites across two continents, whereas the others were single-country studies. In total, 1,681 children from 9 studies were included in our meta-analysis of stool Xpert's diagnostic performance compared to a microbiologic reference standard, and 869 children from 5 studies were included in the comparison against a clinical reference standard. The prevalence of microbiologically confirmed cases per study ranged widely, from 2.6% (14) to 54% (15). The prevalence of clinically confirmed or unconfirmed cases was much higher, ranging from 35% (16) to 100% (17). Table S1 in the supplemental material provides details on clinical reference standard definitions of the included studies.

Studies enrolled children from 0 to 16 years. The ratio of females to males was generally balanced. The percentage of participants with a documented history of TB disease contact, when reported (5/9 studies), ranged from 12% (18) to 56% (19). Most studies did not include information about tuberculin skin test (TST) results. Two studies included only children with HIV (18, 20), and two restricted enrolment to HIV-negative children (16, 21); the remainder had a mixed population.

Table 4.1: Characteristics of included studies. Features of included studies and participants. Studies that included separate comparisons of stool Xpert for microbiological and clinical reference standards have two rows.

Study	Location	No. eligible children	Age in years (range, median [IQR])	No. patients (%)					Clinical features reported	EPTB status, No. EPTB (%)	Reference standard	Samples used for reference standard	Total included in analysis	No. microbiologically confirmed (%)	No. clinically confirmed / unconfirmed cases (%)	No. clinically unlikely TB (%)	No. contaminated cultures (%)
				Females	TB history	TB contact history	TST positive	HIV-positive									
Banada 2016 (15)	South Africa	40	0-15, NR	21/38 (55)	NR	16/38 (42)	NR	16/38 (42)	Cough, EP symptoms, Weight loss	PTB only	**Xpert**	IS, GA	37	20 (54)	-	-	NR
Chipinduro 2017 (17)	Zimbabwe	218	5-16, 10.6 [8-13]	123/218 (56)	17/218 (7.8)	51/218 (23)	NR	111/198 (56)	Cough, Weight loss, Night sweats, Fever, Appetite loss	PTB only[a]	**Culture/ Xpert**	IS	218	19 (8.7)	-	-	NR
											CRS[b]	-	32	-	32 (100)	0 (0)	NR
Hasan 2017 (16)	Pakistan	50	0-15, 6.8 [2-9]	22/50 (44)	NR	27/50 (54)	NR	0/50 (0)	Cough, EP symptoms, Weight loss	PTB only	**Culture/ Xpert**	Sputum, GA	49	11 (22)	-	-	NR
											CRS[b]	-	49	-	17 (35)	32 (65)	NR
Lacourse 2018 (18)	Kenya	165	0-12, 2 [1.1-4.8]	75/165 (45)	NR	20/162 (12)	7/151 (4.6)	165/165 (100)	Cough, Lethargy, Fever, Failure to thrive	PTB only[a]	**Culture/ Xpert**	Sputum, GA	147	11 (7.5)	-	-	NR
											CRS[b]	-	165	-	85 (52)	80 (48)	NR
Marcy 2016 (20)	Burkina Faso, Cambodia, Cameroon, Vietnam	272	0-13, 7.2 [4.1-7.2]	132/272 (49)	49/272 (18)	58/272 (21)	50/272 (18)	272/272 (100)	Cough, Weight loss, Lethargy, Fever, Broad spectrum Abx failure, CXR abnormality	PTB only[a]	**Culture**	GA, IS, NS, string	272	27 (10)	-	-	NR
											CRS[b]	-	272	-	245 (90)	27 (10)	NR

Moussa 2016 (21)	Egypt	115	1-16, NR	45/115 (39)	NR	29/115 (25)	13/67 (19)	0/115 (0)	Cough, Weight loss, Night sweats, Fever, CXR abnormality	PTB only	**Culture**	Sputum, IS	115	36 (31)	-	-	0/115 (0)
Nicol 2013 (22)	South Africa	115	1-15, 2.6 [1.6-4.8]	NR	0/115 (0)	NR	NR	17/115 (15)	Cough, Weight loss, CXR abnormality	PTB only	**Culture**	IS	115	17 (15)	-	-	NR
Orikiriza 2018 (14)	Uganda	357	1-14, NR	178/392 (45)	8/392 (2.0)	76/391 (19)	99/383 (26)	121/388 (31)	Cough, Weight loss, Night sweats, Lethargy, Fever	PTB only[a]	**Culture/ Xpert**	Sputum, IS	349	9 (2.6)	-	-	6/357 (1.7)
Walters 2017 (19)	South Africa	379	0-13, 1.3 [0.8-2.4]	184/379 (49)	27/379 (7.1)	214/379 (56)	82/294 (28)	51/379 (13)	Cough, Weight loss, Fever	Mix of EPTB and PTB, 35/379 (9.2)	**Culture/ Xpert**	GA, IS, NA, string	379	72 (19)	-	-	NR
											CRS[b]	-	351	-	242 (69)	109 (31)	NR

Abx = antibiotics; CRS = clinical reference standard; CXR = chest x-ray; EP = extrapulmonary; EPTB = extrapulmonary TB; GA = gastric aspirate; IQR = interquartile range; IS = induced sputum; NA = nasopharyngeal aspirate; No. = number of; NR = not reported; TST = tuberculin skin test;. [a] =implied only pulmonary TB cases based on collection of respiratory samples only; [b] = definitions of each clinical reference standard are given in Supplementary Table S1

.

4.6.3 Sample processing

Table 4.2 shows the sample preparation steps utilized in each study. In one study (19), two sample preparation methods were attempted, with results ultimately pooled. Most studies (6/9) obtained one stool sample from enrolled children, typically within 24 h of obtaining respiratory samples. Samples were either used immediately or stored for later use, except for one study (20) which used some samples immediately and some after freezing and a second study (19) which stored samples collected at the child's home and immediately used those collected at the health care center. As information on sample storage was not available for all studies, sub-group analysis could not be performed per sample storage method.

The mass of stool utilized, and its collection method, varied: 0.15 g of bulk stool (16), 0.15 g using a sterile loop (17), a flocked rectal swab (22), 0.5 g (21), 0.6g (15), 2 g (20), and 5 g (19). A diluent solution, such as phosphate-buddered saline (PBS), distilled water, or a sucrose solution, was added to the stool before a homogenization, in variable quantities, typically followed by vortexing. Most studies (6/9) reported a period of sample settling before further work-up. Final sample preparation methods were quite varied but included either centrifugation or filtering through syringe filter or gauze, primarily to remove large particles, before final addition of the sample to the Xpert cartridge (
Table 4.2).

Table 4.2: Details of stool sample preparation and processing. Details of stool sample storage and processing for each of the included studies.

Study	No. samples, volume	Stool sample collection timing	Immediate use?	Storage details	Volume stool retrieved	First reagent(s) added to stool	Homogenisation	Specimen settling	Additional reagents and or filtering / processing	Pellet processing	Final sample into cartridge
Banada 2016	1, 5g	NR	No	4°C for 7 days	0.6g	2mL processing buffer (AL buffer, 10% povidone), 2mL Xpert buffer	Vortex with glass beads	30min at RT	All syringe filtered	No pellet	2mL added to cartridge
Chipinduro 2017	1, 5g	Within 24hr of respiratory sample	No	4°C for max 2 days	0.15g using sterile loop	2.4mL PBS	Vortex	20min at RT	1mL supernatant taken, centrifuged at 3200rpm for 15min	Pellet resuspended in 1mL PBS	Diluted 2:1 in buffer, added to cartridge
Hasan 2017	1, NR	Within 24hr of respiratory sample	No	2-8°C for NR days, taken to 3e hospital, stored at -80°C	0.15g	2.4mL PBS	Vortex	20min at RT	1mL supernatant taken, centrifuged at 3500rpm for 15min	Pellet resuspended in 1mL PBS	Diluted 2:1 in buffer, added to cartridge
Lacourse 2018	NR, NR	Within 24hr of respiratory sample	Yes	NA	NR	Equal volume PBS	Manual homogen-isation	12 to 48h at 2-5°C	All filtered through fine filter, vortexed; added to equal volume NaOH-NALC; NRmL PBS added and centrifuged, twice	Pellet resuspended in 0.7mL PBS by vortex	Diluted 2:1 in buffer, added to cartridge
Marcy 2016	NR, 0.5g	NR	Both	Some frozen at NR for NR days	2g	10mL distilled H_2O	Vortex	NR	NRmL supernatant taken, centrifuged at 4000rpm for 20min	Pellet decontam-inated in 10mL 3% NALC-NaOH for 15min at RT; added to 40mL PBS; centrifuged 20min; pellet resuspended in 1mL PBS	Diluted 2:1 in buffer, added to cartridge
Moussa 2016	2, 2g	NR	Yes	NA	0.5g	10mL Sheather's solution (28% sucrose)	Manual homogen-isation, Vortex 30sec	NR	All filtered through funnel gauze; centrifuged at 100g for 1min;	No pellet	0.5mL supernatant, 1.8mL buffer added to cartridge; sit 15min at RT; shake; run

Nicol 2013	1, NR	"At baseline"	No	-80°C within 2hr for max 6 months	0.15g using FLOQ Swabs	2.4mL PBS	Vortex	20min at RT	1mL supernatant taken, centrifuged at 3200rpm for 15min	Pellet resuspended in 1mL PBS	Diluted 2:1 in buffer, added to cartridge
Orikiriza 2018	1, NR	NR	Yes	NA	NR	Saline solution	Vortex	5min at RT	5mL mixture taken, added to NaOH-NALC, vortexed, stand for 20min; PBS added to 50mL and centrifuged at 3000g for 20min at 4°C	Pellet decontaminated with NaOH-NALC method; respun; pellet resuspended in 1.5mL unspecified buffer	0.5mL added to cartridge
Walters 2017	1, 0.3-5g	Within 7 days of respiratory sample	Both	2-8°C for max 3 days if collected at home	<5g	20mL PBS	Vortex	No	5mL mixture taken, added to NALC-NaOH	"Concentration"	Diluted 2:1 in buffer, added to cartridge
	1, 0.3-5g	Within 7 days of respiratory sample	Both	2-8°C for max 3 days if collective at home	1-4g	10mL PBS	Vortex	No	All centrifuged at 3000g at 4°C for 20min	Pellet resuspended in 10mL by vortex for 20sec; centrifuged at 2000g for 1sec, keep supernatant	1mL supernatant added to cartridge

3e = tertiary; max = maximum; NA = not applicable; No. = number of; NR = not reported; PBS = phosphate buffered saline; RT = room temperature; NALC-NaOH = *N*-Acetyl-l-Cysteine–Sodium Hydroxide; tx = anti TB treatment

.

4.6.4 Quality assessment

Figure 4.2 displays the overall risk of bias and applicability concerns of the 9 studies included in our meta-analysis. Figure S1 presents the individual studies' quality assessment results. In the patient selection domain (Figure 4.2), five studies were at low risk of bias, and one study (15) was at high risk of bias due to its use of a case-control design, whereas the remaining eight were either cross-sectional or cohort studies. Risk of bias was high for one study because of convenience sampling (16) and unclear in two studies because of an unclear sampling strategy and the inappropriate exclusions of certain children (17, 21). With respect to applicability, the majority of studies (Table 4.1) included children who presented with symptoms suggestive of TB. Two studies (18, 20) only included children with HIV and, because it is known that Xpert performs differentially in those who are HIV-infected (23), these were scored for applicability concerns as high. One study (15) only tested samples from confirmed TB cases and non-cases, which does not represent a typical clinical scenario, so we also rated applicability concerns as high.

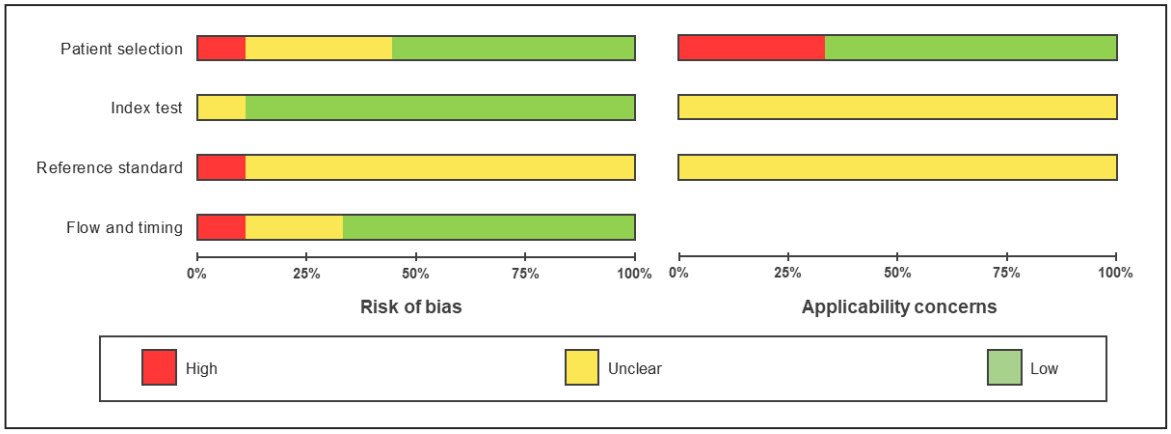


Figure 4.2: QUADAS-2 risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across the 9 included studies The conduct of the index test generally was at low risk of bias, as Xpert is an automated assay with a predefined cut-off of detection that produces a binary response. However, since there is no standardized operating protocol for stool samples and no internationally recommended procedure for sample storage and processing, applicability concerns regarding the index test's conduct are unclear (Figure 4.2).

In light of the inherent limitations of microbiological tests for diagnosing childhood PTB, we classified 8/9 studies as having an unclear risk of bias with respect to correctly classifying the

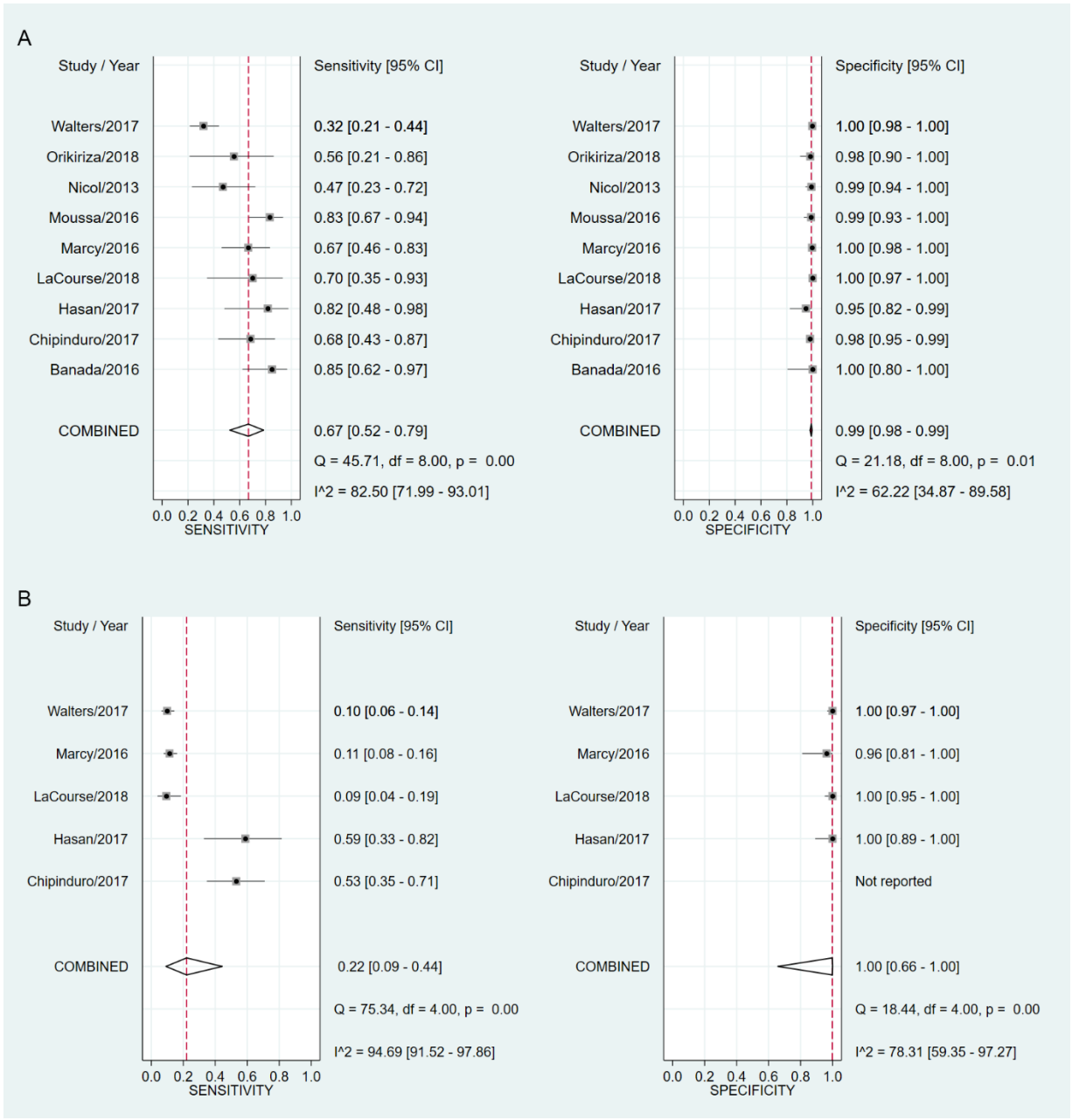
target condition despite having used culture as the reference test. The exception was one study that was scored as having a high risk of bias as its microbiological reference standard did not include culture. Both culture and Xpert are automated assays, so we scored the risk of bias as low regarding test interpretation. Additionally, all studies' reference standards were performed in regional or central reference laboratories, so we expect bias from operator error to be of low concern. Applicability concerns were uniformly unclear.

We scored the risk of bias as low for all studies with respect to the appropriateness of the time interval between index test and reference standard, as all studies reported running stool Xpert within 7 days of specimen collection (Figure 4.2).

4.6.5 Meta-analysis of diagnostic accuracy

For comparison against the microbiological reference standard, sensitivities of stool Xpert varied from 32% (19) to 85% (15), while specificity was uniformly very high (Figure 4.3A). The pooled sensitivity was 67% (95% CI, 52 to 79%) and pooled specificity was 99% (95% CI, 98 to 99%). I^2 values for sensitivity and specificity were 83% (95% CI, 72 to 93%) and 62% (95% CI, 35 to 90%), respectively, indicating high between-study heterogeneity, particularly for sensitivity. For the clinical reference standard comparison, the pooled sensitivity of stool Xpert was 22% (95% CI, 9.0 to 44%), while the specificity was 100% (95% CI, 66 to 100%) (Figure 4.3B).

Although 7/9 studies included children with HIV, only 5/9 studies provided sufficient information to construct two-by-two tables (15, 17, 18, 21, 22) (2 of these studies enrolled only children with HIV (18, 21)) (Figure 4.3C). One study (15) did not provide sufficient information to calculate specificity amongst children with HIV. Data from children that were HIV-negative were available from 5 studies (14, 16, 17, 20, 22) (Figure 4.3D). Using the microbiologic reference standard, among children with HIV, the sensitivity of stool Xpert was 79% (95% CI, 68 to 87%), and the pooled specificity was 99% (95% CI, 94 to 100%) (Figure 4.3C); among those without HIV, the sensitivity was 60% (95% CI, 44 to 79%) and specificity 99% (95% CI, 97 to 100%) (Figure 4.3D). For both sensitivity and specificity, I^2 values were lower in HIV-stratified analyses than when data from all studies were pooled (Table 4.3), suggesting that HIV partially explained the between-study heterogeneity.


A
Study / Year
Sensitivity [95% CI]
Walters/2017
0.32 [0.21 - 0.44]
Orikiriza/2018
0.56 [0.21 - 0.86]
Nicol/2013
0.47 [0.23 - 0.72]
Moussa/2016
0.83 [0.67 - 0.94]
Marcy/2016
0.67 [0.46 - 0.83]
LaCourse/2018
0.70 [0.35 - 0.93]
Hasan/2017
0.82 [0.48 - 0.98]
Chipinduro/2017
0.68 [0.43 - 0.87]
Banada/2016
0.85 [0.62 - 0.97]
COMBINED
0.67 [0.52 - 0.79]
Q = 45.71, df = 8.00, p = 0.00
I^2 = 82.50 [71.99 - 93.01]
0.0 0.2 0.4 0.6 0.8 1.0
SENSITIVITY
Study / Year
Specificity [95% CI]
Walters/2017
1.00 [0.98 - 1.00]
Orikiriza/2018
0.98 [0.90 - 1.00]
Nicol/2013
0.99 [0.94 - 1.00]
Moussa/2016
0.99 [0.93 - 1.00]
Marcy/2016
1.00 [0.98 - 1.00]
LaCourse/2018
1.00 [0.97 - 1.00]
Hasan/2017
0.95 [0.82 - 0.99]
Chipinduro/2017
0.98 [0.95 - 0.99]
Banada/2016
1.00 [0.80 - 1.00]
COMBINED
0.99 [0.98 - 0.99]
Q = 21.18, df = 8.00, p = 0.01
I^2 = 62.22 [34.87 - 89.58]
0.0 0.2 0.4 0.6 0.8 1.0
SPECIFICITY
B
Study / Year
Sensitivity [95% CI]
Walters/2017
0.10 [0.06 - 0.14]
Marcy/2016
0.11 [0.08 - 0.16]
LaCourse/2018
0.09 [0.04 - 0.19]
Hasan/2017
0.59 [0.33 - 0.82]
Chipinduro/2017
0.53 [0.35 - 0.71]
COMBINED
0.22 [0.09 - 0.44]
Q = 75.34, df = 4.00, p = 0.00
I^2 = 94.69 [91.52 - 97.86]
0.0 0.2 0.4 0.6 0.8 1.0
SENSITIVITY
Study / Year
Specificity [95% CI]
Walters/2017
1.00 [0.97 - 1.00]
Marcy/2016
0.96 [0.81 - 1.00]
LaCourse/2018
1.00 [0.95 - 1.00]
Hasan/2017
1.00 [0.89 - 1.00]
Chipinduro/2017
Not reported
COMBINED
1.00 [0.66 - 1.00]
Q = 18.44, df = 4.00, p = 0.00
I^2 = 78.31 [59.35 - 97.27]
0.0 0.2 0.4 0.6 0.8 1.0
SPECIFICITY

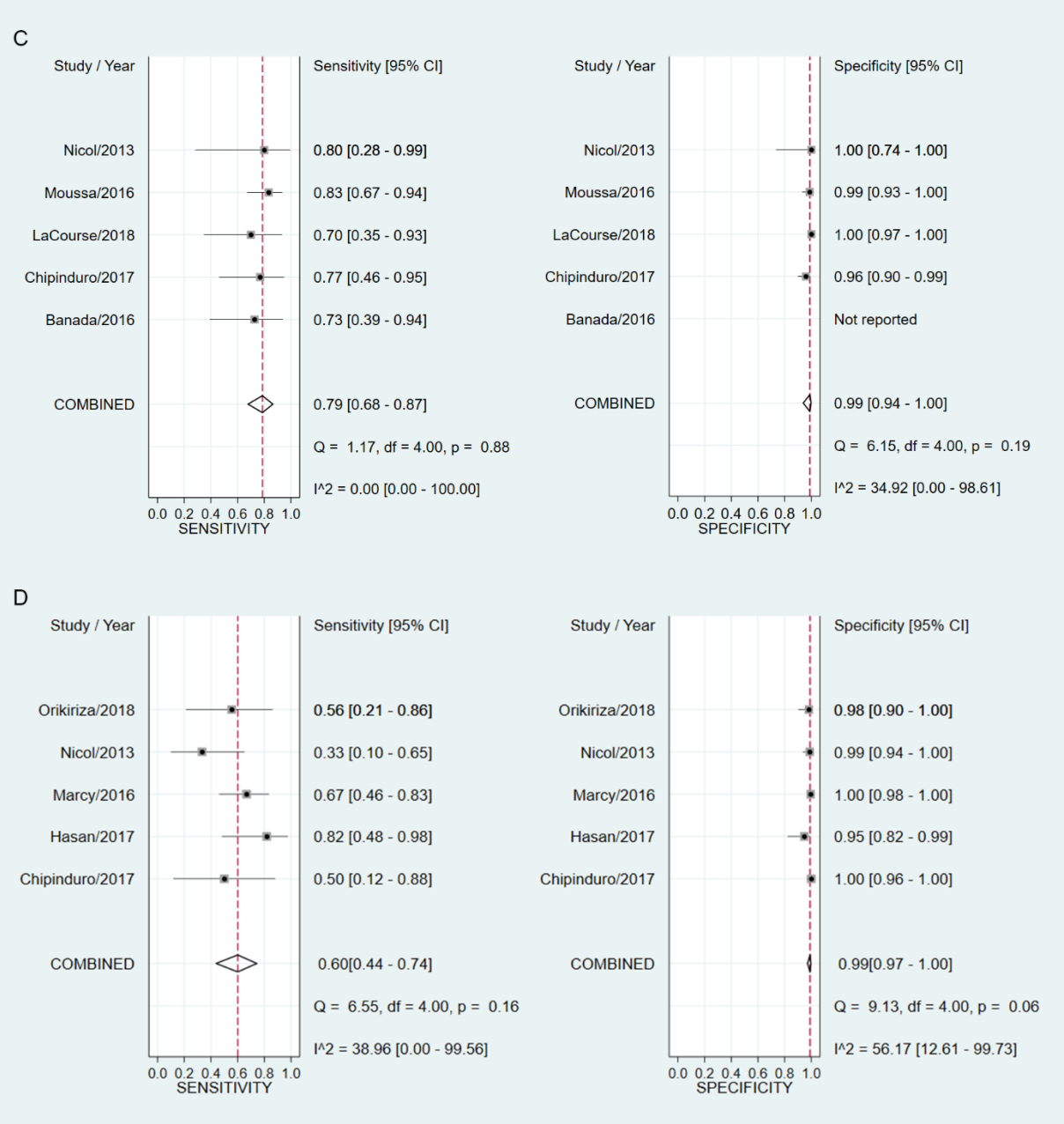


Figure 4.3: Forest plots of stool Xpert's diagnostic performance A: compared to a microbiological reference standard of culture or Xpert positivity on respiratory samples. Two studies (18, 20) presented results from "intention-to-treat" (ITT) analyses, where any child who produced any sample was included, as well as "per protocol" analyses, where only children who produced all requested samples were included. In these instances, we meta-analysed the ITT results to avoid selection bias. B: Forest plots of stool Xpert's diagnostic performance compared to a clinical reference standard of "likely/possibly TB" or "unlikely TB". C: Forest plots of diagnostic performance of stool Xpert in children with HIV compared to a microbiological reference

standard. D: Forest plots of diagnostic performance of stool Xpert in HIV-negative children compared to a microbiological reference standard

Results of the sensitivity analysis in which we excluded the study that did not use mycobacterial culture as part of the reference standard (15) are presented in Fig. S2. Pooled sensitivity and specificity estimates combining data from all studies and data stratified by HIV status were all similar to those estimated in our main analyses, as was between-study heterogeneity. Pooled estimates from our main analysis and from this sensitivity analysis are summarized in Table 4.3.

We undertook two *post hoc* sensitivity analysis. In the first, we sought to determine whether the quantity of stool used for testing was associated with diagnostic accuracy (assuming that a higher mass might increase sensitivity). There were too few studies to estimate pooled accuracy stratified by stool mass used; however, visual inspection of forest plots found no obvious trend to support a minimum quantity (Fig. S3. In the second sensitivity analysis, we evaluated whether the burden of TB in the country where a study was conducted was associated with the accuracy of stool Xpert. As shown in Fig. S4, there was no clear trend to suggest such an association.

Table 4.3: Results of meta-analyses for estimated stool Xpert sensitivity and specificity. I^2-statistic was used to quantify the effect of between study heterogeneity.

	Main results			Sensitivity analysis[a]		
	No. studies included (no. children included)	**Pooled sensitivity (95% CI); I^2 statistic (95% CI)**	**Pooled specificity (95% CI); I^2 statistic (95% CI)**	**No. studies included (no. children included)**	**Pooled sensitivity (95% CI); I^2 statistic (95% CI)**	**Pooled specificity (95% CI); I^2 statistic (95% CI)**
Stool Xpert against microbiological reference standard	9 (14-22) (1681)	67% (52-79); 83 (72-93)	99% (98-99); 62 (35-90)	8 (15-22) (1644)	64% (49-76); 81 (69-93)	99% (98-100); 61 (31-91)
Stool Xpert against clinical reference standard	5 (16-20) (869)	22% (9.0-44); 95 (92-98)	100% (66-100); 78 (59-97)	Not applicable	Not applicable	Not applicable
Stool Xpert against microbiological reference standard in children with HIV	5 (15, 17, 18, 21, 22) (395)	79% (68-87); 0 (0-100)	99% (94-100); 35 (0-99)	5 (17, 18, 21, 22) (379)	80% (68-88); 0 (0-100)	99 (94-100); 51 (0-100)
Stool Xpert against microbiological reference standard in HIV-negative children	7 (14-17, 19, 20, 22) (974)	61% (40-79); 39 (0-100)	99% (98-100); 56 (13-100)	Not applicable	Not applicable	Not applicable

[a]: results from pooled meta-analysis after removal of one study that did not use culture as the microbiological reference standard; CI = confidence interval; no. = number of

4.7 Discussion

In this systematic review and meta-analysis, we found that the sensitivity and specificity of stool Xpert (67% [95% CI, 52 to 79%] and 99% [95% CI, 98 to 99%], respectively) for the diagnosis of microbiologically confirmed childhood PTB were comparable to what has been reported for the performance of Xpert on respiratory specimens (62% [95% credible interval, 51 to 73%] and 98% [95% credible interval, 97 to 99%], respectively) (4). Sensitivity and specificity varied by HIV status. As stool collection is noninvasive, this is of substantial interest for the medical evaluation of children with suspected PTB, but a number of limitations of the existing evidence highlight the need for more research, and greater standardization of testing, before policy formulation.

Among the most important limitations of the evidence base is the lack of data on performance in the subpopulation of children for whom stool Xpert is of greatest potential clinical utility, those under the age of 5 years, and especially the subgroup under the age of 2 years. Only one study compared accuracy between age categories, and a cutoff of 10 years of age was used (17).

We observed substantial between-study heterogeneity in diagnostic accuracy, mostly for sensitivity. Different approaches to participant selection likely contributed to this, in particular the use of a case-control design (15) and nonconsecutive sampling (16, 21), which are at a higher risk of introducing bias into a study. Data also suggested that heterogeneity was partly explained by differences in the prevalence of HIV infection. The higher sensitivity of stool Xpert among children with HIV has also been observed for other specimen types in this population (4, 24), perhaps as a result of more severe TB disease in HIV/TB coinfected children.

We found substantial variability in protocols for performing stool Xpert, with each study taking a unique approach. Differences were seen at all steps: (i) at stool collection, different methods of sampling, numbers of specimens, and volumes of stool were used; (ii) different reagents were added to stool samples before homogenization, and all studies utilized different additional reagents; and (iii) dissimilar filtration methods and decontamination steps were adopted. Future studies should ensure, at minimum, complete reporting of protocols for stool collection processing and testing. A standardized protocol would be of value, as would a standardized stool collection-and-processing kit.

Our systematic review and meta-analysis has a number of strengths. First, all included studies reported using a microbiological reference standard for comparison to stool Xpert, and 8 out of

9 studies used liquid or solid culture. While the imperfect nature of any reference standard for diagnosing pediatric TB means that the true number of affected children is always unknown, the accuracy of stool Xpert against microbiological confirmation is likely a closer estimation of its true accuracy than its performance compared to the clinical reference standard (as symptoms of PTB are non-specific). Second, by systematically assessing each study's sample preparation and processing techniques, we found substantial variability in methods of performing stool Xpert and were also able to identify obstacles to implementation. For example, most protocols required at least one centrifugation step, which is inauspicious in terms of translating this assay to a lower health care system level. Finally, we utilized a sensitive and validated search strategy that covered six languages.

The present work also has some limitations. First, data were insufficient, and there were too few studies for us to perform stratified or meta-regression analyses to assess most demographic-related potential causes of observed heterogeneity. Hence, we suggest that in addition to HIV-stratified results, future studies of stool Xpert should also ensure reporting is stratified by age, gender, and extent of radiographic disease. Second, while we identified wide variability in sampling and stool processing, we could not explore these as sources of heterogeneity or determine if any were potentially superior. Third, we did not include one study concerning stool Xpert on samples from children (25) that was published after our systematic search was completed and therefore was not included in our meta-analysis. However, including it in our pooled analyses did not significantly alter sensitivity or specificity estimates (see Fig. S5 in the supplemental material). Finally, our pooled estimates came from study populations with a high prevalence of T; hence, it is possible that these estimates may not be generalizable to settings of lower TB burdens.

Given that these preliminary studies of stool Xpert suggest high specificity and moderate sensitivity, its potential role in the diagnostic pathway would be as a first-line rule-in test rather than as a triage test to rule out PTB. Studies assessing whether stool Xpert has added value as an add-on test in combination with currently deployed assays will be useful, as will studies assessing the effect of repeat testing on sensitivity.

4.7.1 Conclusion

Preliminary data suggest that the use of Xpert on stool specimens may be potentially useful as a rule-in test, but a standardized stool sample preparation protocol is lacking, and the accuracy

of stool Xpert in children under 5 years old, the subgroup for whom the test could bring the most added value, remains largely unknown.

Acknowledgements

The authors are grateful to Samuel Schumacher for his early critical comments regarding study protocol development.

4.8 Afterward

Since the publication of our systematic review and meta-analysis of stool Xpert in 2019, WHO recommendations regarding stool Xpert have changed, although there is still no manufacturer-recommended stool processing protocol. The 2020 Consolidated Guidelines on Tuberculosis state that Xpert testing should be used as the initial diagnostic test for pediatric TB, as opposed to smear or culture, and stool specimens are listed as an acceptable sample. The certainty of evidence regarding stool Xpert's diagnostic accuracy is low, but the use of stool as a sample for children is a strong recommendation (8). The data we extracted for my study were shared with the authors of the systematic review and meta-analysis that informed these WHO guidelines.

4.9 References

1. World Health Organization. 2018. Global Tuberculosis Report 2018. World Health Organization, Press W, Geneva, Switzerland.
2. Marais BJ, Gie RP, Hesseling AC, Schaaf HS, Lombard C, Enarson DA, Beyers N. 2006. A refined symptom-based approach to diagnose pulmonary tuberculosis in children. Pediatrics 118:e1350-9.
3. World Health Organization. 2013. Policy Update. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. World Health Organization, Geneva, Switzerland.
4. Detjen AK, DiNardo AR, Leyden J, Steingart KR, Menzies D, Schiller I, Dendukuri N, Mandalakas AM. 2015. Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis. Lancet Respir Med 3:451-61.
5. Graham SM, Cuevas LE, Jean-Philippe P, Browning R, Casenghi M, Detjen AK, Gnanashanmugam D, Hesseling AC, Kampmann B, Mandalakas A, Marais BJ, Schito M, Spiegel HM, Starke JR, Worrell C, Zar HJ. 2015. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. Clin Infect Dis 61Suppl 3:S179-87.
6. Cordova J, Shiloh R, Gilman RH, Sheen P, Martin L, Arenas F, Caviedes L, Kawai V, Soto G, Williams DL, Zimic M, Escombe AR, Evans CA. 2010. Evaluation of molecular tools for detection and drug susceptibility testing of Mycobacterium tuberculosis in stool specimens from patients with pulmonary tuberculosis. J Clin Microbiol 48:1820-6.
7. Donald PR, Schaaf HS, Gie RP, Beyers N, Sirgel FA, Venter A. 1996. Stool microscopy and culture to assist the diagnosis of pulmonary tuberculosis in childhood. J Trop Pediatr 42:311-2.
8. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JA, Bossuyt PM. 2011. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529-36.
9. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. 2010. Chapter 10: Analysing and Presenting Results. *In* Collaboration TC (ed), Cochrane Handbok for Systematic Reviews of Diagnostic Test Accuracy, Version 1.0 ed. The Cochrane Collaboration.
10. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. 2005. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 58:982-90.
11. Higgins JP, Thompson SG, Deeks JJ, Altman DG. 2003. Measuring inconsistency in meta-analyses. Bmj 327:557-60.
12. Dwamena B. 2007. MIDAS: Stata module for meta-analytical integration of diagnostic test accuracy studies, vol S456880.
13. Moher D, Liberati A, Tetzlaff J, Altman DG. 2009. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 151:264-9, w64.
14. Orikiriza P, Nansumba M, Nyehangane D, Bastard M, Mugisha IT, Nansera D, Mwanga-Amumpaire J, Boum Y, 2nd, Kumbakumba E, Bonnet M. 2018. Xpert MTB/RIF diagnosis of childhood tuberculosis from sputum and stool samples in a high TB-HIV-prevalent setting. Eur J Clin Microbiol Infect Dis doi:10.1007/s10096-018-3272-0.
15. Banada PP, Naidoo U, Deshpande S, Karim F, Flynn JL, O'Malley M, Jones M, Nanassy O, Jeena P, Alland D. 2016. A novel sample processing method for rapid

detection of tuberculosis in the stool of pediatric patients using the Xpert MTB/RIF assay. PLoS ONE 11 (3) (no pagination).
16. Hasan Z, Shakoor S, Arif F, Mehnaz A, Akber A, Haider M, Kanji A, Hasan R. 2017. Evaluation of Xpert MTB/RIF testing for rapid diagnosis of childhood pulmonary tuberculosis in children by Xpert MTB/RIF testing of stool samples in a low resource setting. BMC Res Notes 10:473.
17. Chipinduro M, Mateveke K, Makamure B, Ferrand RA, Gomo E. 2017. Stool Xpert(R) MTB/RIF test for the diagnosis of childhood pulmonary tuberculosis at primary clinics in Zimbabwe. Int J Tuberc Lung Dis 21:161-166.
18. LaCourse SM, Pavlinac PB, Cranmer LM, Njuguna IN, Mugo C, Gatimu J, Stern J, Walson JL, Maleche-Obimbo E, Oyugi J, Wamalwa D, John-Stewart G. 2018. Stool Xpert MTB/RIF and urine lipoarabinomannan for the diagnosis of tuberculosis in hospitalized HIV-infected children. Aids 32:69-78.
19. Walters E, Van Der Zalm MM, Palmer M, Bosch C, Demers AM, Draper H, Goussard P, Schaaf HS, Friedrich SO, Whitelaw A, Warren R, Gie RP, Hesseling AC. 2017. Xpert MTB/RIF on Stool Is Useful for the Rapid Diagnosis of Tuberculosis in Young Children with Severe Pulmonary Disease. Pediatric Infectious Disease Journal 36:837-843.
20. Marcy O, Ung V, Goyet S, Borand L, Msellati P, Tejiokem M, Nguyen Thi NL, Nacro B, Cheng S, Eyangoh S, Pham TH, Ouedraogo AS, Tarantola A, Godreuil S, Blanche S, Delacourt C. 2016. Performance of Xpert MTB/RIF and Alternative Specimen Collection Methods for the Diagnosis of Tuberculosis in HIV-Infected Children. Clinical Infectious Diseases 62:1161-1168.
21. Moussa H, Bayoumi FS, Mohamed AM. 2016. Gene Xpert for Direct Detection of Mycobacterium Tuberculosis in Stool Specimens from Children with Presumptive Pulmonary Tuberculosis. Ann Clin Lab Sci 46:198-203.
22. Nicol MP, Spiers K, Workman L, Isaacs W, Munro J, Black F, Zemanay W, Zar HJ. 2013. Xpert MTB/RIF testing of stool samples for the diagnosis of pulmonary tuberculosis in children. Clinical Infectious Diseases 57:e18-e21.
23. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. 2014. Xpert(R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. Cochrane Database Syst Rev doi:10.1002/14651858.CD009593.pub3:Cd009593.
24. World Health Organization. 2016. The use of the Xpert MTB/RIF assay for the diagnosis of TB. Meeting Report. World Health Organization, Geneva, Switzerland.
25. Walters E, Scott L, Nabeta P, Demers AM, Reubenson G, Bosch C, David A, van der Zalm M, Havumaki J, Palmer M, Hesseling AC, Ncayiyana J, Stevens W, Alland D, Denkinger C, Banada P. 2018. Molecular detection of Mycobacterium tuberculosis from stool in young children using a novel centrifugation-free processing method. J Clin Microbiol doi:10.1128/jcm.00781-18.

4.10 Supplemental material

Diagnostic accuracy of stool Xpert MTB/RIF for the detection of pulmonary tuberculosis in children: a systematic review and meta-analysis

4.10.1 Supplemental tables

Table S1: Clinical reference standard definitions used by publications included in meta-analysis for "confirmed/unconfirmed" TB.

Study	Clinical reference standards	Notes
Chipinduro 2017 (1)	International guidelines (bacteriological confirmation, symptoms, contact history or immunologic evidence, treatment response) (2)	Chest x-ray not uniformly available to all presenting children
Hasan 2017 (3)	Modified Kenneth-Jones criteria (BCG vaccine and scar, contact history, measles history, chest x-ray, PCM grade 3, immunocompromised/immunosuppressed) (4)	Algorithm utilized by Pakistan Pediatric Association
LaCourse 2018 (5)	International guidelines (bacteriological confirmation, symptoms, chest x-ray, contact history or immunologic evidence, treatment response) (2)	-
Marcy 2016 (6)	International guidelines (bacteriological confirmation, symptoms, chest x-ray, contact history or immunologic evidence, treatment response) (2)	-
Walters 2017 (7)	International guidelines (bacteriological confirmation, symptoms, chest x-ray, contact history or immunologic evidence, treatment response) (2)	Decision to administer treatment not based on research case definitions

Table S2: PRISMA checklist.

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
TITLE / ABSTRACT			
Title	1	Identify the report as a systematic review (+/- meta-analysis) of diagnostic test accuracy (DTA) studies.	1
Abstract	2	Abstract: See PRISMA-DTA for abstracts.	2-3
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Clinical role of index test	D1	State the scientific and clinical background, including the intended use and clinical role of the index test, and if applicable, the rationale for minimally acceptable test accuracy (or minimum difference in accuracy for comparative design).	4
Objectives	4	Provide an explicit statement of question(s) being addressed in terms of participants, index test(s), and target condition(s).	5
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	5
Eligibility criteria	6	Specify study characteristics (participants, setting, index test(s), reference standard(s), target condition(s), and study design) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5-6
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full search strategies for all electronic databases and other sources searched, including any limits used, such that they could be repeated.	Text S1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	6
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Definitions for data extraction	11	Provide definitions used in data extraction and classifications of target condition(s), index test(s), reference standard(s) and other characteristics (e.g. study design, clinical setting).	Text S2
Risk of bias and applicability	12	Describe methods used for assessing risk of bias in individual studies and concerns regarding the applicability to the review question.	6
Diagnostic accuracy measures	13	State the principal diagnostic accuracy measure(s) reported (e.g. sensitivity, specificity) and state the unit of assessment (e.g. per-patient, per-lesion).	7

Synthesis of results	14	Describe methods of handling data, combining results of studies and describing variability between studies. This could include, but is not limited to: a) handling of multiple definitions of target condition. b) handling of multiple thresholds of test positivity, c) handling multiple index test readers, d) handling of indeterminate test results, e) grouping and comparing tests, f) handling of different reference standards	7-8

Adapted From: McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, The PRISMA-DTA Group (2018). Preferred Reporting Items for a Systematic Review and

Section/topic	#	PRISMA-DTA Checklist Item	Reported on page #
Meta-analysis	D2	Report the statistical methods used for meta-analyses, if performed.	7-8
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	7-8
RESULTS			
Study selection	17	Provide numbers of studies screened, assessed for eligibility, included in the review (and included in meta-analysis, if applicable) with reasons for exclusions at each stage, ideally with a flow diagram.	Fig. 4.1
Study characteristics	18	For each included study provide citations and present key characteristics including: a) participant characteristics (presentation, prior testing), b) clinical setting, c) study design, d) target condition definition, e) index test, f) reference standard, g) sample size, h) funding sources	Table 4.1
Risk of bias and applicability	19	Present evaluation of risk of bias and concerns regarding applicability for each study.	Figs 4.2, S1
Results of individual studies	20	For each analysis in each study (e.g. unique combination of index test, reference standard, and positivity threshold) report 2x2 data (TP, FP, FN, TN) with estimates of diagnostic accuracy and confidence intervals, ideally with a forest or receiver operator characteristic (ROC) plot.	Figs 4.3A-D, S2-44
Synthesis of results	21	Describe test accuracy, including variability; if meta-analysis was done, include results and confidence intervals.	11-12
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression; analysis of index test: failure rates, proportion of inconclusive results, adverse events).	11-12 Figs S2-S4
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence.	12-13
Limitations	25	Discuss limitations from included studies (e.g. risk of bias and concerns regarding applicability) and from the review process (e.g. incomplete retrieval of identified research).	13-14
Conclusions	26	Provide a general interpretation of the results in the context of other evidence. Discuss implications for future research and clinical practice (e.g. the intended use and clinical role of the index test).	12-14
FUNDING			
Funding	27	For the systematic review, describe the sources of funding and other support and the role of the funders.	15

Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. JAMA. 2018 Jan 23;319(4):388-396. doi: 10.1001/jama.2017.19163. For more information, visit: **www.prisma-statement.org**. (8).

4.10.2 Supplemental figures

Figure S1: Risk of bias and applicability concerns summary: review authors' judgements about each domain for each of the nine included studies. A "-" indicates high risk of bias; a "?" indicates unclear risk of bias; a "+" indicates low risk of bias.

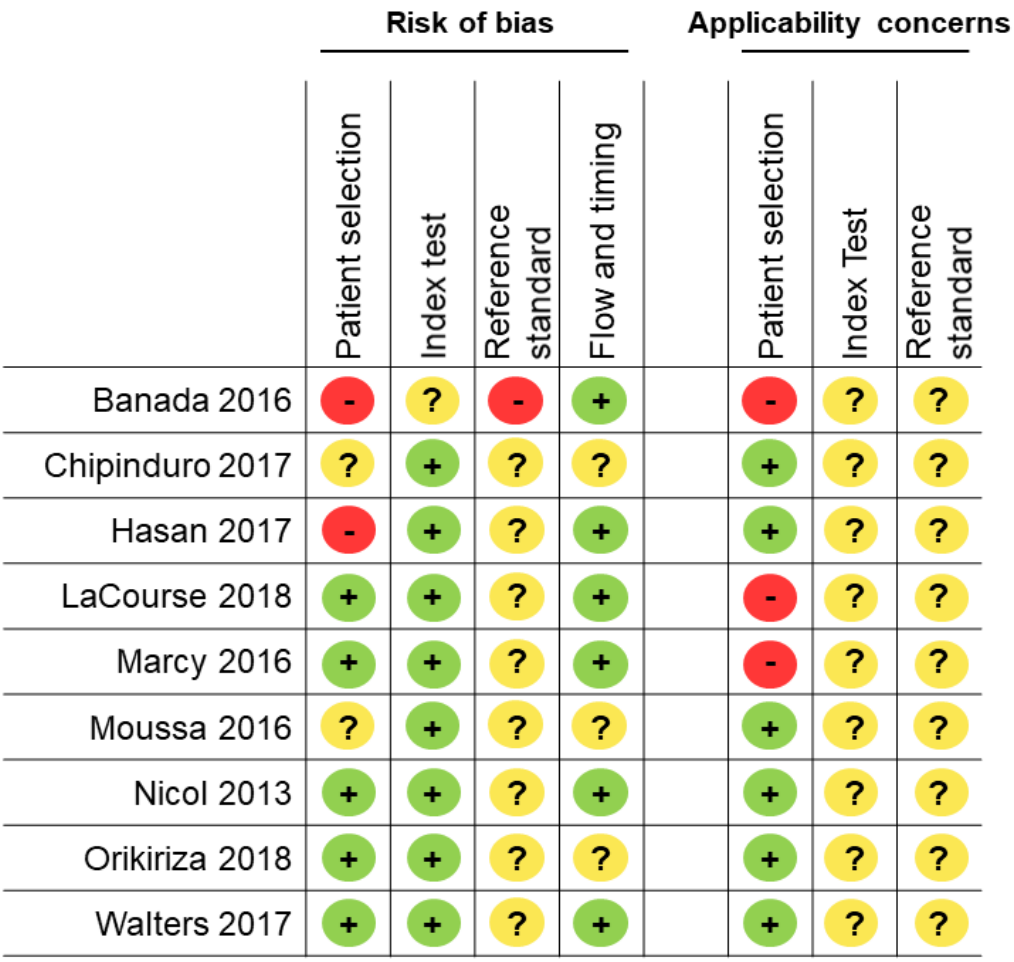


Figure S2a: Forest plots of sensitivity analysis of stool Xpert's diagnostic performance compared to a microbiological reference standard of culture only on respiratory samples using 8 studies; one study (9) that used Xpert as a reference standard was not included here.

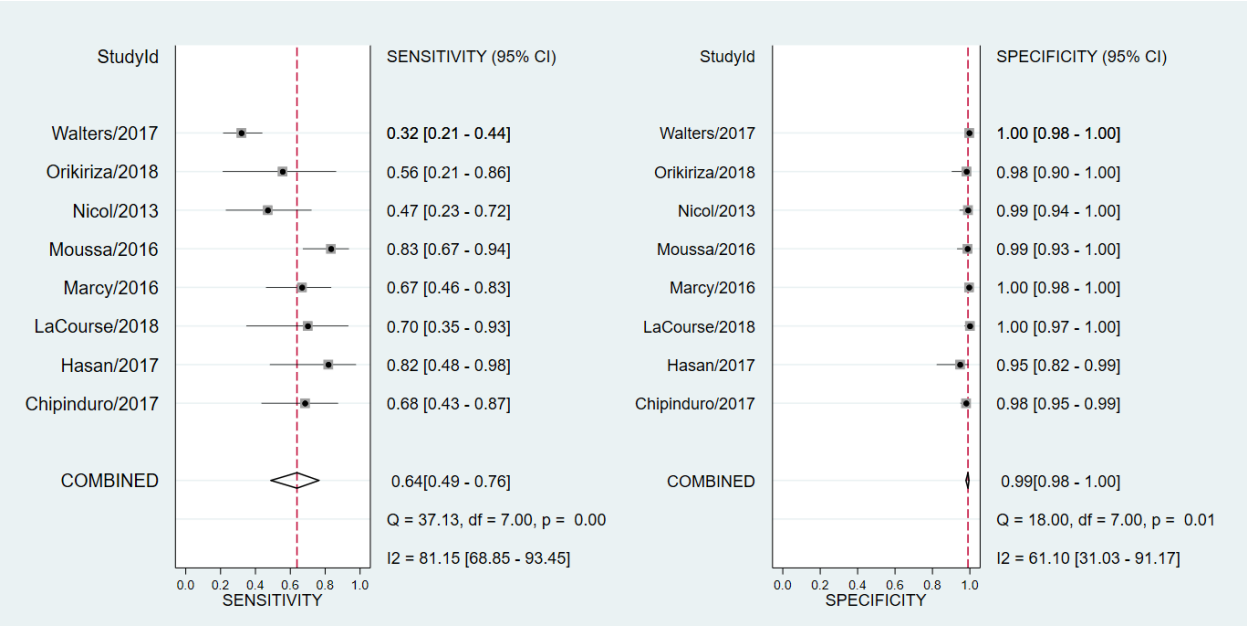


Figure S2b: Forest plots of sensitivity analysis of stool Xpert's diagnostic performance in children living with HIV compared to a microbiological reference standard of culture only on respiratory samples using 8 studies; one study (9) that used Xpert as a reference standard was not included here.

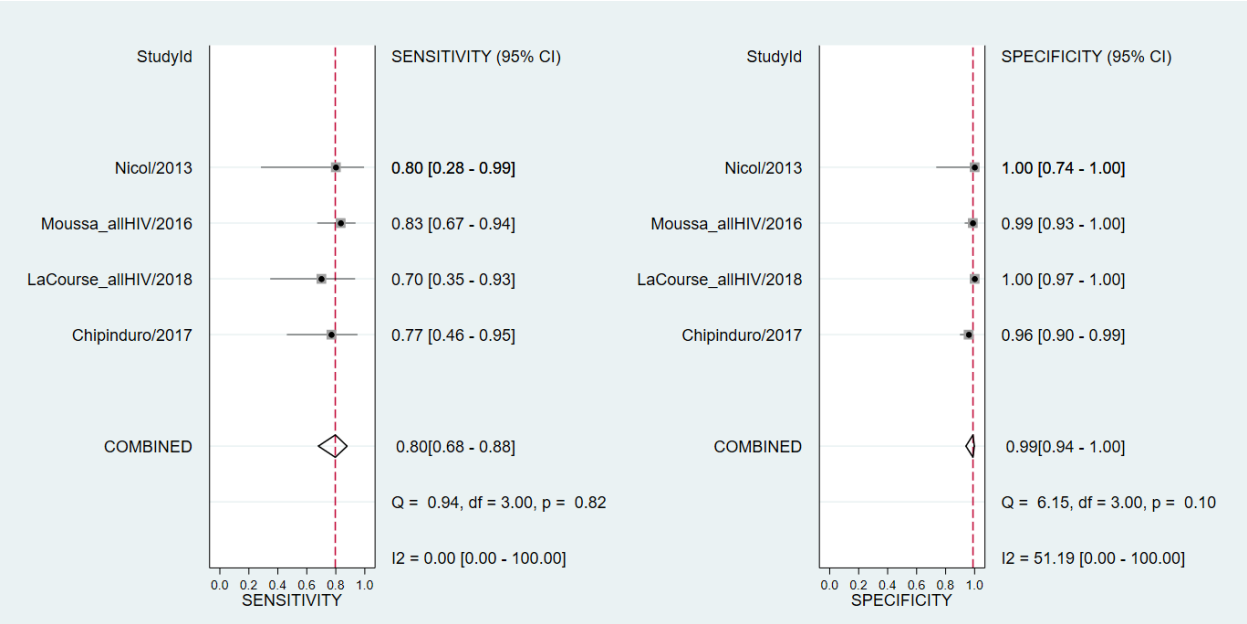


Figure S3: Forest plots of stool Xpert performance arranged by sample mass. The top three listed studies (n=382) utilized 0.15 g of stool collected in bulk, or using a sterile loop or FLOQ swab. The bottom three studies (n=424) used at least 0.5 g of bulk stool.

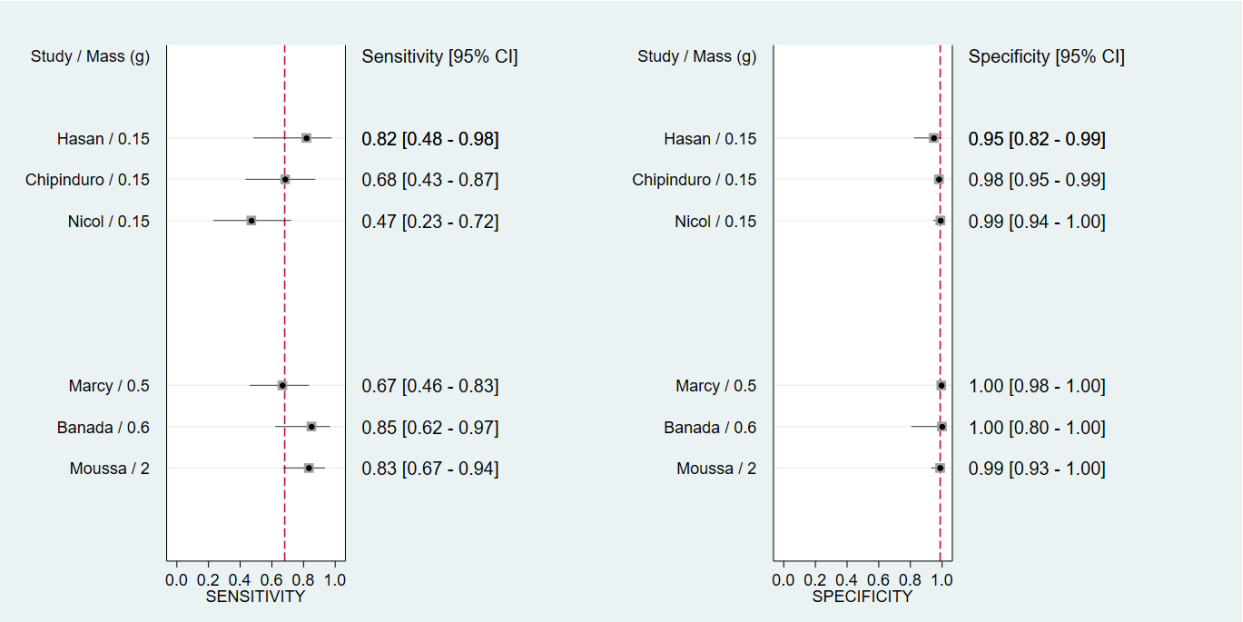


Figure S4: Forest plots of stool Xpert's diagnostic performance compared to a microbiological reference standard, arranged by study countries' TB burdens. Incidence per 100,000 people is presented after the study author. Countries where studies recruited patients are presented in Table 1.

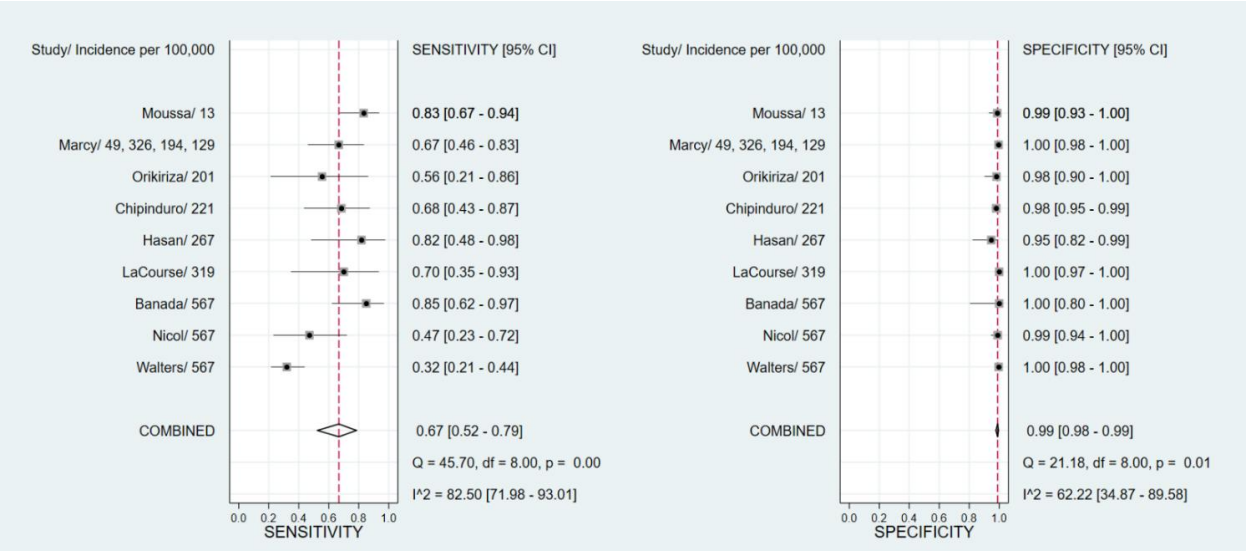
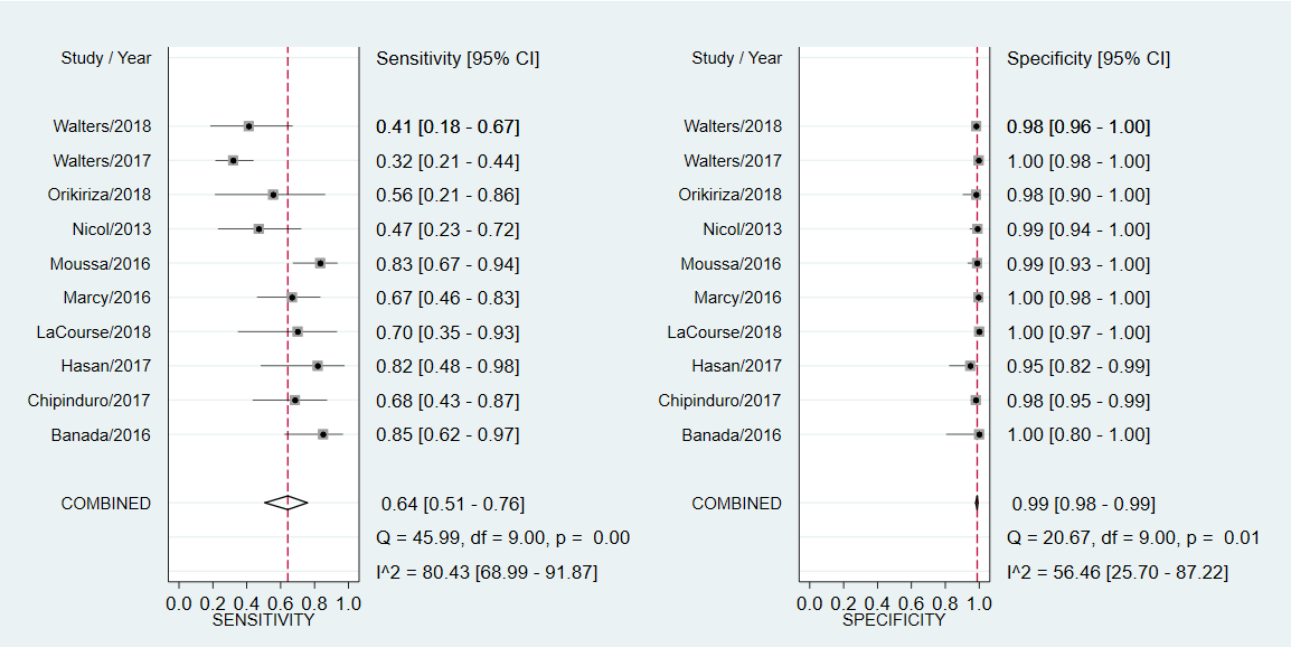


Figure S5: Forest plots of stool Xpert's diagnostic performance compared to a microbiological reference standard of culture or Xpert positivity on respiratory samples using 10 studies, including one study (10) that was published after the date of the systematic search.



4.10.3 Supplemental text

Text S1A: Search strategy used for PubMed

No further filters

(((((((Infan* OR newborn* OR new-born* OR perinat* OR neonat* OR baby OR baby* OR babies OR toddler* OR minors OR minors* OR boy OR boys OR boyfriend OR boyhood OR girl* OR kid OR kids OR child OR child* OR children* OR schoolchild* OR schoolchild OR school child[tiab] OR school child*[tiab] OR adolescen* OR juvenil* OR youth* OR teen* OR under*age* OR pubescen* OR pediatrics[mh] OR pediatric* OR paediatric* OR peadiatric* OR school[tiab] OR school*[tiab] OR prematur* OR preterm*)))) AND ((("tuberculosis"[mesh] OR tuberculos*[tw] OR TB[tw] OR "mycobacterium tuberculosis"[mesh])))) AND ((((("Nucleic Acid Amplification Techniques"[Mesh] OR "molecular diagnostic techniques"[mesh:noexp] OR nucleic acid test*[tw] OR NAAT[tw] OR NAATs[tw] OR NAA[tw] OR molecular assay*[tw] OR molecular diagnos*[tw] OR Molecular technique*[tw] OR Molecular test*[tw] OR polymerase chain reaction*[tw] OR PCR[tw] OR PCRs[tw] OR Xpert[tw] OR GeneXpert[tw] OR cepheid[tw] OR MTB/RIF[tw] OR (nucleic acid[tw] AND amplification[tw]) OR NAT[ti] OR NATs[ti] OR (amplified[tw] AND direct test*[tw])))))) AND ("2008/01/01"[PDat] : "2018/06/15"[PDat]))

Text S1B: Search strategy used for EMBASE

Limit date range to 2008-2018, then combine with AND

No other filters

(Infan* or newborn* or new-born* or perinat* or neonat* or baby or baby* or babies or toddler* or minors or minors* or boy or boys or boyfriend or boyhood or girl* or kid or kids or child or child* or children* or schoolchild* or schoolchild or school child or school child* or adolescen* or juvenil* or youth* or teen* or under*age* or pubescen* or pediatrics or pediatric* or paediatric* or peadiatric* or school or school* or prematur* or preterm*).tw

"Nucleic Acid Amplification Techniques".kw. or nucleic acid test*.tw. or NAAT.tw. or NAATs.tw. or NAA.tw. or molecular assay*.tw. or molecular diagnos*.tw. or molecular technique*.tw. or molecular test*.tw. or polymerase chain reaction*.tw. or PCR.tw. or PCRs*.tw. or Xpert.tw. or GeneXpert.tw. or cepheid.tw. or MTB RIF.tw.

("tuberculosis" or tuberculos* or TB or "mycobacterium tuberculosis").tw.

Text S1C: Search strategy used for Scopus

Limit date range to 2008-2018, then combine with AND

No other filters

(TITLE-ABS-KEY ("tuberculosis" OR tuberculos* OR tb OR "mycobacterium tuberculosis") AND PUBYEAR > 2007 AND PUBYEAR < 2019) AND (TITLE-ABS-KEY ("Nucleic Acid Amplification Techniques" OR nucleic AND acid AND test* OR naat OR naats OR naa OR molecular AND assay* OR molecular AND diagnos* OR molecular AND technique* OR molecular AND test* OR polymerase AND chain AND reaction* OR pcr OR pcrs* OR xpert OR genexpert OR cepheid OR mtb) AND PUBYEAR > 2007 AND PUBYEAR < 2018) AND ((TITLE-ABS-KEY (infan* OR newborn* OR new-born* OR perinat* OR neonat* OR baby* OR babies OR toddler* OR minors* OR boy* OR girl* OR kid* OR child OR child* OR children* OR schoolchild* OR adolescen* OR juvenil* OR youth* OR

teen* OR pubescen* OR pediatrics OR pediatric*) OR TITLE-ABS-KEY (paediatric* OR prematur* OR preterm*)) AND PUBYEAR > 2007 AND PUBYEAR < 2019)

Text S1D: Search strategy used for Cochrane Library

Title, Abstract, Keywords: tuberculosis AND

Search all text: xpert OR NAAT AND

Search all text: child* OR pediat*

Text S2: Data extracted from each of the 9 studies included in the meta analysis

(i) bibliographic information;

(ii) index test: number of samples tested, place of sample testing, clinical setting, unit of analysis, number of indeterminate results;

(iii) reference standard: reference standard employed and remarks, number of contaminated cultures;

(iv) participant information: descriptive study population information, age demographic, sex,

total number of participants, negative population, different populations included in

study, HIV status, smear status, EPTB status, other comorbidities, disease contacts, history of TB, symptoms;

(v) study information: author-defined study design, study timing, sampling strategy,

study location, study time period, place of sample testing, study location, study time

period;

(vi) diagnostic performance data: for (a) microbiological and (b) clinical reference standards, by HIV-status or smear-status or EPTB status if information was available: numbers of true positives, true negatives, false positives, and false negatives, sensitivity and confidence intervals, specificity and confidence intervals, positive and negative predictive values, number of TB cases, number of reference standard negative controls;

(vii) stool processing information;

4.10.4 Supplemental references

1. Chipinduro M, Mateveke K, Makamure B, Ferrand RA, Gomo E. 2017. Stool Xpert(R) MTB/RIF test for the diagnosis of childhood pulmonary tuberculosis at primary clinics in Zimbabwe. Int J Tuberc Lung Dis 21:161-166.
2. Graham SM, Cuevas LE, Jean-Philippe P, Browning R, Casenghi M, Detjen AK, Gnanashanmugam D, Hesseling AC, Kampmann B, Mandalakas A, Marais BJ, Schito M, Spiegel HM, Starke JR, Worrell C, Zar HJ. 2015. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. Clin Infect Dis 61Suppl 3:S179-87.
3. Hasan Z, Shakoor S, Arif F, Mehnaz A, Akber A, Haider M, Kanji A, Hasan R. 2017. Evaluation of Xpert MTB/RIF testing for rapid diagnosis of childhood pulmonary tuberculosis in children by Xpert MTB/RIF testing of stool samples in a low resource setting. BMC Res Notes 10:473.
4. National Tuberculosis Control Program. 2015. National guidelines for the management of tuberculosis in Pakistan. Coordination MoNHSR, Pakistan, Islamabad, Pakistan.
5. LaCourse SM, Pavlinac PB, Cranmer LM, Njuguna IN, Mugo C, Gatimu J, Stern J, Walson JL, Maleche-Obimbo E, Oyugi J, Wamalwa D, John-Stewart G. 2018. Stool Xpert MTB/RIF and urine lipoarabinomannan for the diagnosis of tuberculosis in hospitalized HIV-infected children. Aids 32:69-78.
6. Marcy O, Ung V, Goyet S, Borand L, Msellati P, Tejiokem M, Nguyen Thi NL, Nacro B, Cheng S, Eyangoh S, Pham TH, Ouedraogo AS, Tarantola A, Godreuil S, Blanche S, Delacourt C. 2016. Performance of Xpert MTB/RIF and Alternative Specimen Collection Methods for the Diagnosis of Tuberculosis in HIV-Infected Children. Clinical Infectious Diseases 62:1161-1168.
7. Walters E, Van Der Zalm MM, Palmer M, Bosch C, Demers AM, Draper H, Goussard P, Schaaf HS, Friedrich SO, Whitelaw A, Warren R, Gie RP, Hesseling AC. 2017. Xpert MTB/RIF on Stool Is Useful for the Rapid Diagnosis of Tuberculosis in Young Children with Severe Pulmonary Disease. Pediatric Infectious Disease Journal 36:837-843.
8. McInnes MDF, Moher D, Thombs BD, McGrath TA, Bossuyt PM, Clifford T, Cohen JF, Deeks JJ, Gatsonis C, Hooft L, Hunt HA, Hyde CJ, Korevaar DA, Leeflang MMG, Macaskill P, Reitsma JB, Rodin R, Rutjes AWS, Salameh JP, Stevens A, Takwoingi Y, Tonelli M, Weeks L, Whiting P, Willis BH. 2018. Preferred Reporting Items for a Systematic Review and Meta-analysis of Diagnostic Test Accuracy Studies: The PRISMA-DTA Statement. Jama 319:388-396.
9. Banada PP, Naidoo U, Deshpande S, Karim F, Flynn JL, O'Malley M, Jones M, Nanassy O, Jeena P, Alland D. 2016. A novel sample processing method for rapid detection of tuberculosis in the stool of pediatric patients using the Xpert MTB/RIF assay. PLoS ONE 11 (3) (no pagination).
10. Walters E, Scott L, Nabeta P, Demers AM, Reubenson G, Bosch C, David A, van der Zalm M, Havumaki J, Palmer M, Hesseling AC, Ncayiyana J, Stevens W, Alland D, Denkinger C, Banada P. 2018. Molecular detection of Mycobacterium tuberculosis from stool in young children using a novel centrifugation-free processing method. J Clin Microbiol doi:10.1128/jcm.00781-18.

5 CHAPTER 5: MANUSCRIPT 3: BAYESIAN LATENT CLASS ANALYSIS PRODUCED DIAGNOSTIC ACCURACY ESTIMATES THAT WERE MORE INTERPRETABLE THAN COMPOSITE REFERENCE STANDARDS FOR EXTRAPULMONARY TUBERCULOSIS TESTS

5.1 Preface

Another patient group in whom TB may be challenging to diagnose are people with extrapulmonary TB. Although extrapulmonary forms of TB comprise a substantial fraction of the global TB burden, the development of diagnostics for this diverse group has lagged. Due to the lack of a gold standard, the performance of existing assays for diagnosing extrapulmonary TB is not very well-characterised, which makes evaluating the accuracy of newer molecular tests difficult. Other analytical approaches that can account for the tests' imperfect accuracies would be appropriate, such as Bayesian latent class analysis. Having less-biased estimates of test's diagnostic accuracy can allow for improved estimates of disease burden, as well as better resource planning and allocation.

In this manuscript, I have used Bayesian latent class analysis to estimate the diagnostic accuracy of five tests, including Xpert MTB/RIF, for three forms of extrapulmonary TB: TB lymphadenitis, TB meningitis, and TB pleuritis. The results are also compared to previously published estimates derived from a series of composite reference standards.

This manuscript has been accepted for publication at *Diagnostic and Prognostic Research*.

5.2 Title page

Bayesian latent class analysis produced diagnostic accuracy estimates that were more interpretable than composite reference standards for extrapulmonary tuberculosis tests

Emily MacLean[a,b], Mikashmi Kohli[a,b], Lisa Köppel[c], Ian Schiller[e], Surendra K. Sharma[d], Madhukar Pai[a,b], Claudia M. Denkinger[a,c], Nandini Dendukuri[a,e]

[a] McGill International TB Centre, Research Institute of the McGill University Health Centre; Montréal, Canada

[b] Department of Epidemiology, Biostatistics and Occupational Health, McGill University; Montréal, Canada

[c] Division of Tropical Medicine, Center of Infectious Diseases, University of Heidelberg; Heidelberg, Germany

[d] Department of Medicine, All India Institute of Medical Sciences; New Delhi, India

[e] Department of Medicine, McGill University Health Centre; Montréal, Canada

5.3 Abstract

Background

Evaluating the accuracy of extrapulmonary tuberculosis (TB) tests is challenging due to lack of a gold standard. Latent class analysis (LCA), a statistical modeling approach, can adjust for reference tests' imperfect accuracies to produce less biased test accuracy estimates than those produced by commonly used methods like composite reference standards (CRSs). Our objective is to illustrate how Bayesian LCA can address the problem of an unavailable gold standard and demonstrate how it compares to using CRSs for extrapulmonary TB tests.

Methods

We re-analysed a dataset of presumptive extrapulmonary TB cases in New Delhi, India for three forms of extrapulmonary TB. Results were available for culture, smear microscopy, Xpert MTB/RIF, and a non-microbiological test, cytopathology/histopathology or adenosine deaminase (ADA). A diagram was used to define assumed relationships between observed tests and underlying latent variables in the Bayesian LCA with input from an inter-disciplinary team. We compared the results to estimates obtained from a sequence of CRSs defined by increasing numbers of positive reference tests necessary for positive disease status.

Results

Data were available from 298, 388, and 230 individuals with presumptive TB lymphadenitis, meningitis, and pleuritis, respectively. Using Bayesian LCA, estimates were obtained for accuracy of all tests and for extrapulmonary TB prevalence. Xpert sensitivity neared that of culture for TB lymphadenitis and meningitis but was lower for TB pleuritis, and specificities of all microbiological tests approached 100%. Non-microbiological tests' sensitivities were high, but specificities were only moderate, preventing disease rule-in. CRSs' only provided estimates of Xpert and these varied widely per CRS definition. Accuracy of the CRSs also varied by definition, and no CRS was 100% accurate.

Conclusion

Unlike CRSs, Bayesian LCA takes into account assumptions about test performance resulting in accuracy estimates that are easier to interpret. LCA should receive greater consideration for evaluating extrapulmonary TB diagnostic tests.

5.4 Background

Extrapulmonary tuberculosis (TB) comprised approximately 16% of the global TB burden in 2019, or 1.6 million cases (1). This estimate is highly uncertain as a reliable 'gold standard' to diagnose extrapulmonary TB cases is unavailable. The requisite non-respiratory samples are difficult to obtain, and existing diagnostic tests are not optimized for these typically paucibacillary samples. Mycobacterial culture and sputum smear microscopy, the conventional microbiological tests for TB, cannot detect bacteria at low counts, although cultures are substantially more sensitive than smears. Similarly, the limit of detection of Xpert MTB/RIF (Xpert) (Cepheid, USA), the World Health Organization-endorsed molecular test, is too high to detect TB in samples with low numbers of bacteria (2). Consequently, multiple microbiological and clinical tests often requiring invasive techniques are relied on to make a diagnosis, with each test's accuracy varying by extrapulmonary specimen type (3). Therefore, when evaluating the performance of a new extrapulmonary TB test, the conventional tests cannot be treated as perfect reference standards as this will lead to bias (4).

In response, a composite reference standard (CRS) that combines results from multiple tests and clinical assessments in some pre-defined way is often employed to classify individuals as extrapulmonary TB-positive or -negative (5-8). Though CRSs are designed with the goal of improving upon the accuracy of the individual component tests, it is recognized that they are imperfect themselves and, moreover, they have been criticized for making sub-optimal use of gathered data (9, 10). Specifically, the most commonly used CRSs ignore the sensitivity and specificity of individual assessment components (i.e., tests and clinical symptoms) and treat them all as having similar accuracy. CRSs also assume that the component tests are independent of each other. However, it is possible that different imperfect tests are conditionally dependent, meaning that multiple tests may be more likely to be simultaneously falsely negative or falsely positive than if they were independent. For example, given all microbiological tests for extrapulmonary TB rely on bacterial load, they may all produce false negative results for a paucibacillary diseased individual (10).

Latent class analysis (LCA) is a statistical modeling solution to address these issues (11). LCA can model the accuracy of each imperfect diagnostic test, as well as dependence between tests, to simultaneously estimate disease prevalence and the sensitivity and specificity of all tests at hand (10, 12). Bayesian LCA can further include reliable prior information on disease prevalence or test accuracy parameters, when available, e.g., high specificity values for microbiological tests. It has been applied to estimate diagnostic test accuracy for latent TB

infection (13), childhood pulmonary TB (14), chlamydia (15), and *Helicobacter pylori* infection (16), among other diseases.

Extrapulmonary TB has traditionally received less attention than pulmonary TB owing to its less infectious nature. Nonetheless, it presents a significant burden to healthcare systems and patients, particularly due to the difficulty in diagnosis, and the risk of severe outcomes with certain extrapulmonary TB forms is high (e.g., TB meningitis and miliary TB). LCA may overcome limitations of naïve methods to produce better prevalence and test accuracy estimates, better approximating the true burden. However, it must be acknowledged that LCA methods are considered methodologically complex and difficult to validate (11).

Therefore, to illustrate the steps involved in conducting an LCA and compare with CRS, we re-analysed an existing dataset of extrapulmonary samples from adults with presumptive extrapulmonary TB (17). The original study evaluated Xpert accuracy using a series of CRSs for TB meningitis, TB lymphadenitis, and TB pleuritis. Resultant estimates of Xpert sensitivity, specificity, and disease prevalence varied widely depending on the CRS used (17), making them difficult to interpret. Our objective was to use Bayesian LCA to estimate the diagnostic accuracy of all the available tests and to discuss the advantages and challenges of this approach.

5.5 Methods

5.5.1 Primary dataset details

The primary dataset comprised all extrapulmonary samples from adults with presumptive extrapulmonary TB received by the All-India Institute of Medical Sciences (AIIMS), a tertiary hospital in New Delhi, India in 2012 (17). No participants had taken anti-TB therapy (ATT) for longer than two weeks. All samples underwent testing with Xpert, an automated PCR-based assay that detects mycobacterial-specific DNA; BACTEC Mycobacteria Growth Indicator Tube (Becton Dickinson, USA) liquid culture or Lowenstein-Jensen solid culture; and Ziehl-Neelsen (acid-fast bacilli) sputum smear microscopy. As conventional pulmonary TB tests perform sub-optimally for extrapulmonary TB, non-specific assays are also deployed in conjunction with TB testing (18). Regarding non-microbiological tests, for presumptive TB lymphadenitis, results were available from cytopathology/histopathology, wherein local cells or tissues are examined for pathological patterns such as caseating necrosis (19), while for presumptive TB meningitis and pleuritis (solid tissue or fluid), levels of deaminase (ADA), an enzyme expressed in

leukocytes associated with granulatomous reactions (20), were available. Additionally, for each participant, type of extrapulmonary sample tested (indicating extrapulmonary TB form), resistance to rifampicin, initiation of ATT, and treatment response were available. Each participant had one result per test. Demographic covariates and clinical symptoms were unavailable. As this was a secondary analysis of previously collected data which had received ethical approvals and informed consent from all participants, ethical approval was not necessary.

In the original publication (17), the authors focused on the then-novel Xpert test, reporting its sensitivity and specificity against culture. Recognizing culture's imperfect performance, they subsequently compared Xpert to a series of CRSs, resulting in multiple estimates of Xpert sensitivity and specificity.

5.5.2 Latent class model specification

Diagrammatic representation: We first created heuristic diagrams for each extrapulmonary TB form (meningitis, lymphadenitis, pleuritis) to illustrate our assumptions about the relationships between observed test results and unobservable (latent) extrapulmonary TB status (14) (Figure 5.1). These diagrams identify the measurand of each test, i.e., the quantity it measures. Culture, smear, and Xpert use different techniques to measure the presence of *Mycobacterium tuberculosis* in the extrapulmonary sample. We assumed that this was equivalent to their measurand being extrapulmonary TB itself. In the case of TB meningitis and pleuritis, the ADA test was also deployed. The ADA test result is determined by the latent variable 'ADA level' and not by extrapulmonary TB. Similarly, in the situation of TB lymphadenitis, the result on cytopathology/histopathology is determined by 'change in cell morphology' rather than extrapulmonary TB. The nonspecific tests, ADA test and cytopathology/histopathology, do not measure the target condition extrapulmonary TB per se; instead, their measurands are signals (e.g., inflammation) caused by extrapulmonary TB or other diseases. Therefore, for each of the LCA models we assume there are four possible latent classes resulting from combinations of the latent measurands: 1. Extrapulmonary TB-positive, non-specific measurand-positive; 2. Extrapulmonary TB-positive, non-specific measurand-negative; 3. Extrapulmonary TB-negative, non-specific measurand-positive; and 4. Extrapulmonary TB-negative, non-specific measurand-negative. This approach differs from the widely used two-class LCA model which assumes all observed tests measure the target condition, i.e., extrapulmonary TB in our applications. We preferred the four-class LCA as it leads to greater interpretability of the latent classes. The parameters resulting from a two-class LCA (prevalence of extrapulmonary TB and accuracy of

tests with respect to extrapulmonary TB) may be obtained easily as a subset of the four-class LCA.

Observation with smear microscopy or cytopathology/histopathology, bacterial growth on culture, and detection of mycobacterial DNA using Xpert are all increasingly likely as bacterial burden increases. Contrastingly, in people with paucibacillary extrapulmonary TB, these tests will tend to show negative results. We thus hypothesised that the underlying bacterial burden may create dependence between test results, i.e. dependence between tests conditional in the first two latent classes of extrapulmonary TB positive subjects (21), even though tests are based on different mechanisms (Figure 5.1).

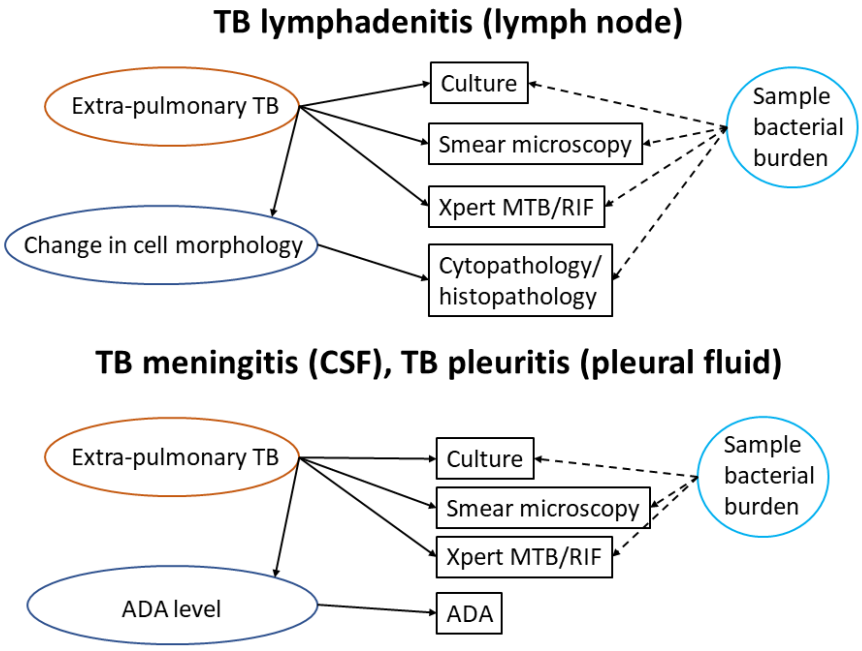


Figure 5.1: Heuristic model. Heuristic model. The model shows the assumed relationships between latent classes (ovals), diagnostic test results (rectangles), and random effect representing sample bacterial burden (circle). ADA – adenosine deaminase; CSF – cerebral spin fluid; TB – tuberculosis*Statistical model*: The observed diagnostic test results were assumed to be a mixture of results from the four underlying latent classes. The unknown parameters of the model were the prevalence of the four latent classes and each test's sensitivity and specificity with respect to its measurand. Using these parameters, we can further determine the prevalence of each measurand, the accuracy of the non-specific measurand and the accuracy of the non-specific test for extrapulmonary TB. For example, the prevalence of extrapulmonary TB can be obtained by adding the prevalence of the two classes with extrapulmonary TB, with or without the non-specific condition. We introduced a random effect, corresponding to sample

bacterial burden, to account for the conditional dependence among microbiological tests and cytopathology/histopathology in the group of people with extrapulmonary TB. In people without extrapulmonary TB, all test outcomes were considered conditionally independent. Individuals with invalid or missing test results were assumed to be missing at random and retained in the analysis, with missing test results imputed by Bayesian imputation.

5.5.3 Bayesian model estimation

Using a four-class rather than a two-class approach increases the number of unknown parameters in the model and increases concerns for non-identifiability, i.e., the lack of a unique solution to the model. By constraining each test's accuracy parameters to be determined only by its measurand, we limited the number of parameters added, resulting in fewer parameters to be estimated than available degrees of freedom for all three forms of extrapulmonary TB (see Supplemental methods for details on identifiability).

We used a Bayesian approach to fit the latent class models (see Supplement for model likelihoods and prior distributions). As the posterior distributions of the parameters of interest (sensitivity, specificity, prevalence) could not be computed analytically, we sampled from the posterior distributions using a Markov Chain Monte Carlo (MCMC) approach with the rjags package (Version 4-8) through Rstudio (Version 3.5.2). Non-informative priors were used for all models with truncated prior distributions for the non-specific tests' sensitivities and specificities to contain them above 50% and avoid label switching (mirror solutions) (see Supplement). The Supplement contains details of model specifications and sampling, and further programs for data preparation and model checking are available in a repository: https://osf.io/9wdb3/?view_only=730fb3e7d9114405bb51075748703054.

5.5.4 LCA model validation

There is no ideal way to validate the results of an LCA due to the lack of a perfect reference test. As in previous work (14, 15) we used an indirect approach. For each test pattern, we compared the observed frequency of receiving ATT with the LCA-derived probability of disease. If the LCA was valid, we expected to observe that as the LCA-derived probability of TB increased, the probability of ATT would also increase.

5.5.5 Composite reference standards

We used the same definitions for the series of CRSs as the original publication (17). "CRS 1+" was defined as any one component test of the CRS being positive versus all four tests being

negative; “CRS2+” was defined as any two component tests of the CRS being positive versus all four tests being negative; and so on to CRS 4+. For each extrapulmonary TB form, the latent class model was used to estimate the sensitivity and specificity of each CRS. Each individual’s probability of disease, as computed by LCA, was used as the reference standard. The Supplement contains the expressions and code used for these computations.

5.6 Results

5.6.1 Dataset description

The original study had 1376 samples. After excluding specimens for five forms of extrapulmonary TB not considered in the current analyses, there remained 299 lymph node samples for TB lymphadenitis, 388 pleural fluid samples for TB pleuritis, and 230 cerebral spinal fluid (CSF) samples for TB meningitis.

5.6.2 Bayesian latent class analysis: estimated prevalence and diagnostic accuracies

Estimates from the latent class analysis (median values and 95% credible intervals (CrI)) of sensitivity and specificity for each diagnostic test are shown in Table 5.1, organized by form of extrapulmonary TB. Prevalence of TB lymphadenitis, TB meningitis, and TB pleuritis were estimated as 60.5% (95% CrI, 54.1 to 67.9), 15.9% (95% CrI, 9.70 to 24.3) and 35.3% (95% CrI, 26.7 to 48.8), respectively. Note that these are not population-level estimates, but rather prevalence estimates among recruited participants at a tertiary care facility. Supplementary Table S1 shows the probabilities of each of the four latent classes.

Test performance varied by specimen type, as previously observed (3, 22-24). Culture and Xpert sensitivities were highest for TB lymphadenitis, and lower for TB meningitis and TB pleuritis. Sensitivity of ADA and cytopathology/histopathology with respect to extrapulmonary TB were high, but imperfect specificities prevented disease rule-in. As expected, sensitivity and specificity point estimates of these tests were higher with respect to their measurands than with respect to the target condition, extrapulmonary TB (Table 5.2). Specificities for culture, Xpert, and smear were universally almost perfect.

5.6.3 Composite reference standard analysis

As in the earlier publication (17), we confirmed that the four composite reference standards provided four estimates of Xpert accuracy for each form extrapulmonary TB. Regardless of the extrapulmonary TB form, when disease-positivity was defined by the presence of any one

positive test result, CRS 1+ classified most individuals as disease positive and therefore had the highest sensitivity and lowest specificity (Table 5.3). This was observed across extrapulmonary TB forms. Correspondingly, with increasingly stringent CRS definition of disease-positivity, the sensitivities declined across all disease forms, while specificities increased, as the number of false positives decreased.

Table 5.1: Bayesian latent class analysis-derived diagnostic accuracies of tests for each form of extrapulmonary TB

	Culture	Xpert	Smear microscopy	Cytopathology / Histopathology	ADA
TB lymphadenitis					
Sensitivity (%) (95% CrI)	90.1 (80.3, 95.4)	86.6 (77.2, 92.2)	26.8 (20.6, 33.7)	98.7 (96.1, 99.7)	NA
Specificity (%) (95% CrI)	99.3 (96.2,100)	98.5 (94.4, 100)	99.4 (97.0, 100)	83.5 (74.6, 93.2)	NA
TB meningitis					
Sensitivity (%) (95% CrI)	60.5 (43.2, 82.7)	52.6 (36.2, 73.3)	27.5 (14.9, 42.6)	NA	83.1 (64.8, 94.5)
Specificity (%) (95% CrI)	99.2 (96.8, 100)	99.5 (97.6, 100)	98.6 (96.3, 99.7)	NA	90.7 (83.8, 98.1)
TB pleuritis					
Sensitivity (%) (95% CrI)	75.4 (56.1, 94.5)	37.7 (27.2, 49.8)	15.4 (9.4, 23.4)	NA	94.6 (88.8, 98.1)
Specificity (%) (95% CrI)	99.4 (97.3, 100)	96.9 (93.8, 99.0)	99.3 (97.7, 99.9)	NA	74.7 (65.1, 90.3)

Lymph node samples were used for TB lymphadenitis; CSF samples were used for TB meningitis; pleural fluid samples were used for TB pleuritis. The performance of all tests in this table are the estimates with respect to the target condition, extrapulmonary TB.

ADA – adenosine deaminase; CrI – credible interval; CSF – cerebral spinal fluid; NA – not applicable; TB – tuberculosis.

Table 5.2: Cytopathology/histopathology and ADA test performance with respect to each extrapulmonary TB form and measurands LCA

	Cytopathology / Histopathology with respect to EPTB form	ADA test with respect to EPTB form	Cytopathology /histopathology with respect to measurand	ADA test with respect to measurand
TB lymphadenitis				
Sensitivity (%) (95% CrI)	98.7 (96.1, 99.7)	NA	99.8 (98.2, 100)	NA
Specificity (%) (95% CrI)	83.5 (74.6, 93.2)	NA	91.7 (79.6, 99.5)	NA
TB meningitis				
Sensitivity (%) (95% CrI)	NA	83.1 (64.8, 94.5)	NA	91.9 (73.3, 99.6)
Specificity (%) (95% CrI)	NA	90.7 (83.8, 98.1)	NA	95.0 (86.5, 99.8)
TB pleuritis				
Sensitivity (%) (95% CrI)	NA	94.6 (88.8, 98.1)	NA	97.5 (91.9, 98.1)
Specificity (%) (95% CrI)	NA	74.7 (65.1, 90.3)	NA	86.8 (70.1, 99.4)

The performance of each test with respect to type of EPTB, otherwise referred to as the target condition, and with respect to the measurand is provided. Regarding non-specific tests ADA and cytopathology/histopathology, we have discerned between their performance at measuring their measurands versus the target condition, EPTB. This paramaterisation more accurately captures the nuance of the testing scenario, as 'target condition' and 'measurand' are not interchangeable entities(31). Reassuringly, they both performed better at measuring their measurand (ADA level or change in cell morphology) than measuring EPTB.

ADA – adenosine deaminase; CrI – credible interval; EPTB – extrapulmonary TB; NA – not applicable; TB – tuberculosis.

Table 5.3: Diagnostic accuracy of a series of composite reference standards for each form of extrapulmonary TB

	CRS 1+	CRS 2+	CRS 3+	CRS4+
TB lymphadenitis				
Sensitivity (%) (95% CI)	99.8 (97.0, 100)	97.1 (89.1, 99.5)	81.8 (74.8, 84.2)	26.4 (24.2, 27.2)
Specificity (%) (95% CI)	83.9 (80.9, 92.3)	99.6 (98.4, 100)	100 (100, 100)	100 (100, 100)
TB meningitis				
Sensitivity (%) (95% CI)	95.0 (84.6, 99.6)	64.4 (46.3, 87.1)	44.4 (31.3, 61.9)	17.8 (12.5, 24.8)
Specificity (%) (95% CI)	88.3 (84.8, 94.1)	98.8 (97.8, 99.9)	100 (100, 100)	100 (100, 100)
TB pleuritis				
Sensitivity (%) (95% CI)	98.9 (96.0, 99.9)	75.5 (56.3, 94.8)	31.1 (23.1, 39.2)	9.00 (6.70, 11.4)
Specificity (%) (95% CI)	73.7 (67.2, 88.0)	99.3 (97.9, 100)	100 (99.8, 100)	100 (100, 100)

CRS 1+ indicates any one positive test in CRS versus all four parameters being negative. CRS 2+: any two tests positive in CRS versus all four parameters being negative. CRS 3+: any three tests positive in CRS versus all four parameters being negative. CRS 4+: all four tests positive in CRS versus all four parameters being negative(17). Lymph node samples were used for TB lymphadenitis; CSF samples were used for TB meningitis; pleural fluid samples were used for TB pleuritis.

CI – confidence interval; CRS – composite reference standard; CSF – cerebral spinal fluid; TB – tuberculosis.

5.6.4 Model fit

As shown in Table 5.4, Table 5.5, and Table 5.6, expected counts for all observed test result patterns generally resembled observed frequencies, indicating that the data did not deviate from the proposed model. There was no evidence that residual correlation deviated significantly from 0, implying there was no evidence of unaccounted conditional dependence; see Supplemental results, Figure S1, and Figure S2 for further details on model fit.

5.6.5 Probability of extrapulmonary TB and association with probability of receiving ATT

Table 5.4, Table 5.5, and Table 5.6 also display the LCA-derived probability of each extrapulmonary TB form for an individual with a particular test result pattern. In all three examples, many patterns were associated with high probability of extrapulmonary TB (close to 1) or a low probability (close to 0). The most difficult subjects to classify were those with a positive result only on the non-specific test. Contrastingly, using the CRSs, all subjects would be classified as diseased or non-diseased with 100% probability, regardless of the CRSs' performance. Consider that for TB pleuritis (Table 5.6), individuals with positive results for Xpert and ADA but negative results for culture and smear would be classified as disease-positive by CRS1+ and CRS2+ but disease-negative by CRS3+ and CRS4+. Using LCA, their probability of having TB pleuritis was estimated as 0.54, reflecting the lack of certainty in their true classification based on the available evidence.

The observed probability of ATT was usually 100% whenever at least one test produced a positive result. This was true even when the calculated probability of extrapulmonary TB is low, an observation that was consistent across disease forms. For example, in CSF, with three microbiological tests negative but ADA positive, the probability of TB meningitis was only 33%, but the probability of ATT was 96% (26/27 patients). In pleural fluid, with all tests negative except Xpert, the probability of TB pleuritis was 1.6%, but all four patients with this test pattern received ATT. Thus, we did not find the probability of receiving ATT to be informative about the validity of the LCA models.

Table 5.4: Observed counts, expected counts, and TB lymphadenitis probability by test result pattern

Culture	Xpert	Smear	Cytopathology/ histopathology	No. observed	No. expected by LCA (95% CrI)	Probability of TB lymphadenitis by LCA (95% CrI)	Probability of TB lymphadenitis by CRS			
							CRS1+	CRS2+	CRS3+	CRS4+
-	-	-	-	96	92 (85, 98)	0.003 (0, 0.08)	0	0	0	0
-	-	-	+	21	22 (15, 29)	0.18 (0.04, 0.73)	1	0	0	0
-	+	-	-	1	2 (0, 6)	0.18 (0.01, 0.89)	1	0	0	0
-	+	-	+	10	9 (5, 15)	0.98 (0.85, 0.99)	1	1	0	0
+	-	-	+	15	14 (8, 20)	0.99 (0.94, 0.99)	1	1	0	0
+	-	+	+	2	2 (1, 5)	1.00 (0.99, 1)	1	1	1	0
+	+	-	+	85	86 (75, 98)	1.00 (0.99, 1)	1	1	1	0
+	+	+	-	1	0 (0, 1)	1.00 (0.99, 1)	1	1	1	0
+	+	+	+	42	40 (31, 49)	1.00 (0.99, 1)	1	1	1	1
NA	-	-	-	1	NE	0.002 (0, 0.01)	NE	0	0	0
NA	-	-	+	1	NE	0.47 (0.28, 0.85)	1	NE	0	0
NA	+	-	+	13	NE	1 (0.99, 1)	1	1	NE	0
NA	+	+	+	4	NE	1 (0.99, 1)	1	1	1	NE
NA	NA	-	NA	7	NE	0.53 (0.45, 0.61)	NE	NE	NE	0

Tests were all performed in lymph node samples (n=299). A "+" indicates positive test result and "-" indicates negative test result. CRS 1+ indicates any one positive test in CRS versus all four parameters being negative. CRS 2+: any two tests positive in CRS versus all four parameters being negative. CRS 3+: any three tests positive in CRS versus all four parameters being negative. CRS 4+: all four tests positive in CRS versus all four parameters being negative.

CrI – credible interval; LCA – latent class analysis; NA – value not available. NE – result not estimable. No. – number of; Smear – smear microscopy; TB – tuberculosis.

Table 5.5: Observed counts, expected counts, and TB meningitis probability by test result pattern

Culture	Xpert	Smear	ADA	No. observed	No. expected by LCA (95% CrI)	Probability of TB meningitis by LCA (95% CrI)	Probability of TB meningitis by CRS			
							CRS1+	CRS2+	CRS3+	CRS4+
-	-	-	-	170	166 (158,173)	0.01 (0, 0.04)	0	0	0	0
-	-	-	+	27	29 (22,37)	0.33 (0.08, 0.76)	1	0	0	0
-	-	+	-	2	2 (1,6)	0.004 (0, 0.07)	1	0	0	0
-	+	-	+	3	2 (1,4)	0.96 (0.06, 0.99)	1	1	0	0
+	-	-	-	1	2 (1,6)	0.30 (0.03, 0.93)	1	0	0	0
+	-	-	+	6	4 (2,7)	0.97 (0.08, 1)	1	1	0	0
+	+	-	+	7	6 (4,10)	1.0 (0.99, 1)	1	1	1	0
+	+	+	-	2	1 (0,3)	1.0 (0.99, 1)	1	1	1	0
+	+	+	+	6	6 (4,9)	1.0 (0.99, 1)	1	1	1	1
-	-	-	NA	6	NE	0.03 (0, 0.08)	NE	0	0	0

Tests were all performed in CSF samples (n=230). A "+" indicates positive test result and "-" indicates negative test result. CRS 1+ indicates any one positive test in CRS versus all four parameters being negative. CRS 2+: any two tests positive in CRS versus all four parameters being negative. CRS 3+: any three tests positive in CRS versus all four parameters being negative. CRS 4+: all four tests positive in CRS versus all four parameters being negative.

ADA – adenosine deaminase; CrI – credible interval; CSF – cerebral spinal fluid; LCA – latent class analysis; NA – value not available. NE – result not estimable. No. – number of; Smear – smear microscopy; TB – tuberculosis.

Table 5.6: Observed counts, expected counts, and TB pleuritis probability by test result pattern

Culture	Xpert	Smear	ADA	No. observed	No. expected by LCA (95% CrI)	Probability of TB pleuritis by LCA (95% CrI)	Probability of TB pleuritis by CRS			
							CRS1+	CRS2+	CRS3+	CRS4+
-	-	-	-	180	175 (162, 185)	0.008 (0, 0.03)	0	0	0	0
-	-	-	+	84	86 (75, 98)	0.33 (0.06, 0.78)	1	0	0	0
-	-	+	-	1	1 (0, 4)	0.005 (0, 0.07)	1	0	0	0
-	+	-	-	4	6 (2, 11)	0.016 (0, 0.10)	1	0	0	0
-	+	-	+	4	4 (2, 7)	0.54 (0.05, 0.92)	1	1	0	0
+	-	-	-	1	4 (1, 7)	0.70 (0.2, 0.99)	1	0	0	0
+	-	-	+	50	45 (36, 53)	0.99 (0.96, 1)	1	1	0	0
+	-	+	+	3	3 (1, 6)	1.0 (0.99, 1)	1	1	1	0
+	+	-	-	2	1 (0, 3)	0.98 (0.81, 0.99)	1	1	0	0
+	+	-	+	24	26 (20, 33)	1.0 (0.99, 1)	1	1	1	0
+	+	+	+	11	12 (8, 16)	1.0 (0.99, 1)	1	1	1	1
NA	-	-	-	3	NE	0.02 (0.01, 0.05)	NE	0	1	0
NA	-	-	+	7	NE	0.53 (0.33, 0.84)	1	NE	0	0
NA	+	-	-	1	NE	0.29 (0.10, 0.62)	1	NE	0	0
NA	+	-	+	6	NE	0.98 (0.96, 1)	1	1	NE	0
NA	+	+	-	1	NE	0.59 (0.28, 0.93)	1	1	NE	0
NA	+	+	+	4	NE	0.99 (0.98, 1)	1	1	1	NE
NA	+	+	NA	1	NE	0.99 (0.96, 1)	1	1	NE	NE
NA	NA	-	NA	1	NE	0.32 (0.23, 0.45)	NE	NE	NE	0

Tests were all performed in pleural fluid samples (n=388). A "+" indicates positive test result and "-" indicates negative test result. CRS 1+ indicates any one positive test in CRS versus all four parameters being negative. CRS 2+: any two tests positive in CRS versus all

four parameters being negative. CRS 3+: any three tests positive in CRS versus all four parameters being negative. CRS 4+: all four tests positive in CRS versus all four parameters being negative.

ADA – adenosine deaminase; CrI – credible interval; LCA – latent class analysis; NA – value not available; NE – result not estimable. No. – number of; Smear – smear microscopy; TB – tuberculosis.

5.7 Discussion

Producing correct estimates of diagnostic test accuracy is challenging without a perfect reference standard. We used Bayesian LCA to estimate multiple tests' accuracies for three forms of extrapulmonary TB, along with disease prevalence. We also estimated the accuracy of a series of CRSs for these same conditions. By employing these two methods, we hope to demonstrate the utility of Bayesian LCA for evaluating diagnostic test accuracy. We observed that each test's sensitivity varied by extrapulmonary TB form, and none was perfect. Specificities were generally very high with respect to the target condition, except for ADA and cytopathology/histopathology. Culture sensitivity, often treated as 100%, was imperfect and only slightly higher than Xpert's for TB lymphadenitis and meningitis; indeed, India's Index-TB Guidelines recommend Xpert for these two disease forms (18). Though there is no way to validate these models, the results were in-keeping with our expectations. For example, for all forms of extrapulmonary TB we found that the sensitivity of culture was greater than the sensitivity of Xpert which, in turn, was greater than the sensitivity of smear, as would be expected based on the knowledge of the mechanisms behind these tests, while all of them had near perfect specificity. The non-specific tests always had higher accuracy with respect to the measurand they were designed to measure than with respect to extrapulmonary TB.

We calculated that no CRS was 100% accurate; rather, accuracy varied depending on the rule by which the CRS was defined. Further, there is no way of knowing which rule provided the true measure of disease status. With extrapulmonary TB, the diagnostic tests that comprise the CRSs are themselves imperfect, so assuming 100% sensitivity and specificity is unreasonable and ignores relevant test-specific information. Such use would result in biased index test accuracy estimates, with the true values obscure (10). When using LCA, there is no assumption that reference tests perform perfectly. Instead, LCA incorporates all available tests results, concurrently adjusting for their unique accuracies and between-test dependence: this more comprehensive approach more closely approximates the real-world setting where each test brings a different quantum of information by the target condition. In doing so, LCA produces one figure that, based on stipulated assumptions, can be interpreted as the best estimate; for example, the latent class model estimated Xpert sensitivity for TB meningitis as 53% (95% CrI: 36-73). Contrast this with the original study, where the series of CRSs resulted in four different Xpert sensitivity estimates for TB meningitis, 33%, 50%, 70%, and 100% (17). It is impossible for the reader to know which was the true measure of test accuracy. In this way, LCA-derived

values are more clinically interpretable, as the reader does not need to discern between a series of values and select one as the best estimate.

In our study we constructed a four-class latent class model as we felt it was more likely to achieve the desired separation into ‘extrapulmonary TB’ and ‘not extrapulmonary TB’. Depending on the combination of observed test results, a two-class model may have resulted in the combining of two classes where either extrapulmonary TB or the non-specific measurand was present into one class, potentially leading to biased estimates of test accuracy (25). We did fit the two-class latent model as well, for comparison (Tables S2 and S3). For TB lymphadenitis and TB meningitis, results from the two-class and four-class models were very similar because the two discordant classes had relatively low prevalence (Table S1). In the case of TB pleuritis, the two-class model gave a prevalence estimate that appeared equivalent to the probability of the latent classes where either extrapulmonary TB or ADA was positive (Table S2 and S3), resulting in slightly lower point estimates of the sensitivities of the microbiological tests and a slightly lower specificity of the ADA test.

We relied on a multi-disciplinary team of experts when creating our model, as goodness-of-fit metrics may fail to indicate model misspecification (26). Consider that pathological signs on cytopathology/histopathology may be attributable to causes other than TB lymphadenitis. This means that individuals who had positive cytopathology/histopathology signals would be a mix of people with extrapulmonary TB and people with some other disease. Choosing a 4-class model allowed us to distinguish between conditions that could produce positive test results and prevented grouping a mix of extrapulmonary TB-positive and -negative people together in our ‘diseased’ condition.

We found that the LCA-derived probability of extrapulmonary TB was not a good predictor for receiving ATT, as even in cases of low disease probability, patients received treatment. Seemingly, one, and certainly two, positive test results was sufficient to commence ATT. This is perhaps not surprising given the high TB prevalence in the study setting and the very high mortality risk of, for example, TB meningitis (27), so clinicians would rather ‘treat now and ask questions later’. When diagnosing extrapulmonary TB, clinicians also consider clinical variables that were unavailable in our dataset (18). It is worth emphasizing that obtaining diagnostic test accuracy estimates is unlike making a clinical decision. Here, we have constructed a model to estimate test performance and have attempted to be transparent in unknowns, assumptions, and subjective choices, but other parameterisations are certainly possible.

5.7.1 Strengths

We estimated the prevalence of three forms of extrapulmonary TB; understanding prevalence in a particular healthcare setting is critical to planning and policy making. Using LCA, we have made the best possible use of data by incorporating results from all available tests to determine sensitivities and specificities, while adjusting for the possibility of between-test dependence. Unlike with CRSs, we did not ignore any observed test results. Consider that for CRS1+, a single positive test result defines disease positivity; if there are three other negative test results, those three are functionally non-informative. Obtaining specimens for most forms of extrapulmonary TB is invasive and requires trained healthcare workers, so ensuring collected data are used to their best potential is an ethical decision.

5.7.2 Limitations

First, as with any statistical model, the latent class models we have fit cannot be shown to be the true models. However, our models were reasonably well-specified, as evidenced by good agreement between observed and expected test result patterns and low residual correlation between test results. In some applications, external information, such as the proportion of patients with a given test pattern who were treated, provides useful information to validate the model. Here, this information was not very informative due to the sparse nature of the datasets. Second, LCA has been characterised as ‘black box-y’ (28) and cautions have been raised that model misspecification is difficult to detect (26). Certainly its mechanisms are less intuitive to understand than Boolean decision rules like those often used when defining CRSs, but the theory underpinning LCA is well-defined and transparent (29, 30). We have attempted to be clear by providing DAGs illustrating our assumptions of the relationships within the model. A final, general limitation is that parameter estimates depend on the available data. The available datasets had a small sample size and did not contain demographic or clinical assessment variables, resulting in poor precision of the estimates. Additionally, this prevented any relevant subgroup analyses.

5.7.3 Conclusion

Basic methods like two-by-two table calculations and CRSs are known to produce imperfect estimates of diagnostic test accuracy. Latent class analysis, which can reflect knowledge of the individual tests used for diagnosis, should receive greater consideration in evaluating new tests’ performance.

5.8 Declarations

5.8.1 Ethics approval and consent to participate

As this was a secondary analysis of previously collected data which had received ethical approvals and informed consent from all participants, ethical approval was not necessary.

5.8.2 Consent for publication

Not applicable.

5.8.3 Availability of data and materials

The datasets analysed during the current study are available in the Open Science Framework repository, https://osf.io/9wdb3/?view_only=730fb3e7d9114405bb51075748703054.

5.8.4 Competing interests

None.

5.8.5 Funding

This work was supported by the Fonds de Recherche du Québec – Santé [to EM]; Canadian Institutes of Health Research Project Grant [PJT-156039 to ND]; the German Center for Infection Research, site Heidelberg, Germany [to LK and CMD]; and Canada Research Chair award [241486 to MP]. None of the funding sources had any involvement in the study.

5.8.6 Acknowledgements

None.

5.9 References

1. World Health Organization. Global Tuberculosis Report 2020. Geneva: World Health Organization; 2020.
2. Chakravorty S, Simmons AM, Rowneki M, Parmar H, Cao Y, Ryan J, et al. The New Xpert MTB/RIF Ultra: Improving Detection of Mycobacterium tuberculosis and Resistance to Rifampin in an Assay Suitable for Point-of-Care Testing. mBio. 2017;8(4).
3. Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, et al. Xpert((R)) MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. The Cochrane database of systematic reviews. 2018;8:Cd012768.
4. Schiller I, van Smeden M, Hadgu A, Libman M, Reitsma JB, Dendukuri N. Bias due to composite reference standards in diagnostic accuracy studies. Statistics in medicine. 2016;35(9):1454-70.
5. Christopher DJ, Schumacher SG, Michael JS, Luo R, Balamugesh T, Duraikannan P, et al. Performance of Xpert MTB/RIF on pleural tissue for the diagnosis of pleural tuberculosis. The European respiratory journal. 2013;42(5):1427-9.
6. Lusiba JK, Nakiyingi L, Kirenga BJ, Kiragga A, Lukande R, Nsereko M, et al. Evaluation of Cepheid's Xpert MTB/Rif test on pleural fluid in the diagnosis of pleural tuberculosis in a high prevalence HIV/TB setting. PLoS One. 2014;9(7):e102702.
7. Meldau R, Peter J, Theron G, Calligaro G, Allwood B, Symons G, et al. Comparison of same day diagnostic tools including Gene Xpert and unstimulated IFN-γ for the evaluation of pleural tuberculosis: a prospective cohort study. BMC pulmonary medicine. 2014;14:58.
8. Wen H, Li P, Ma H, Lv G. Diagnostic accuracy of Xpert MTB/RIF assay for musculoskeletal tuberculosis: a meta-analysis. Infection and drug resistance. 2017;10:299-305.
9. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. Journal of clinical epidemiology. 2009;62(8):797-806.
10. Dendukuri N, Schiller I, de Groot J, Libman M, Moons K, Reitsma J, et al. Concerns about composite reference standards in diagnostic research. Bmj. 2018;360:j5779.
11. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard--a systematic review. Am J Epidemiol. 2014;179(4):423-31.
12. Yang I, Becker MP. Latent variable modeling of diagnostic accuracy. Biometrics. 1997;53(3):948-58.
13. Pai M, Dendukuri N, Wang L, Joshi R, Kalantri S, Rieder HL. Improving the estimation of tuberculosis infection prevalence using T-cell-based assay and mixture models. Int J Tuberc Lung Dis. 2008;12(8):895-902.
14. Schumacher SG, van Smeden M, Dendukuri N, Joseph L, Nicol MP, Pai M, et al. Diagnostic Test Accuracy in Childhood Pulmonary Tuberculosis: A Bayesian Latent Class Analysis. Am J Epidemiol. 2016;184(9):690-700.
15. Dendukuri N, Hadgu A, Wang L. Modeling conditional dependence between diagnostic tests: a multiple latent variable model. Statistics in medicine. 2009;28(3):441-61.
16. Jekarl DW, Choi H, Kim JY, Lee S, Gweon TG, Lee HK, et al. Evaluating Diagnostic Tests for Helicobacter pylori Infection Without a Reference Standard: Use of Latent Class Analysis. Annals of laboratory medicine. 2020;40(1):68-71.
17. Sharma SK, Kohli M, Chaubey J, Yadav RN, Sharma R, Singh BK, et al. Evaluation of Xpert MTB/RIF assay performance in diagnosing extrapulmonary tuberculosis among adults in a tertiary care centre in India. The European respiratory journal. 2014;44(4):1090-3.
18. Central TB Division - Ministry of Health and Family Welfare; Government of India. Index-TB Guidelines: Guidelines on extra-pulmonary tuberculosis for India. New Delhi: All India Institute of Medical Sciences; 2016.

19. Jain D, Ghosh S, Teixeira L, Mukhopadhyay S. Pathology of pulmonary tuberculosis and non-tuberculous mycobacterial lung disease: Facts, misconceptions, and practical tips for pathologists. Seminars in diagnostic pathology. 2017;34(6):518-29.
20. Michot JM, Madec Y, Bulifon S, Thorette-Tcherniak C, Fortineau N, Noël N, et al. Adenosine deaminase is a useful biomarker to diagnose pleural tuberculosis in low to medium prevalence settings. Diagnostic microbiology and infectious disease. 2016;84(3):215-20.
21. Qu Y, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. Biometrics. 1996;52(3):797-810.
22. Kim YW, Kwak N, Seong MW, Kim EC, Yoo CG, Kim YW, et al. Accuracy of the Xpert® MTB/RIF assay for the diagnosis of extra-pulmonary tuberculosis in South Korea. Int J Tuberc Lung Dis. 2015;19(1):81-6.
23. Jing H, Lu Z, Deng Y, Gao D, Li L, Graviss E, et al. Evaluation of Xpert MTB/RIF in detection of pulmonary and extrapulmonary tuberculosis cases in China. International Journal of Clinical and Experimental Pathology. 2017;10:4847-51.
24. Hillemann D, Rüsch-Gerdes S, Boehme C, Richter E. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. Journal of clinical microbiology. 2011;49(4):1202-5.
25. Schofield MR, Maze MJ, Crump JA, Rubach MP, Galloway R, Sharples KJ. On the robustness of latent class models for diagnostic testing with no gold standard. Statistics in medicine. 2021;40(22):4751-63.
26. Albert PS, Dodd LE. A cautionary note on the robustness of latent class models for estimating diagnostic error without a gold standard. Biometrics. 2004;60(2):427-35.
27. Seddon JA, Tugume L, Solomons R, Prasad K, Bahr NC. The current global situation for tuberculous meningitis: epidemiology, diagnostics, treatment and outcomes. Wellcome Open Res. 2019;4:167.
28. Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. 1 ed. Oxford: Oxford University Press; 2003.
29. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika. 1974;61(2):215-31.
30. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am J Epidemiol. 1995;141(3):263-72.
31. Bossuyt PM. Testing COVID-19 tests faces methodological challenges. Journal of clinical epidemiology. 2020;126:172-6.

5.10 Supplemental material

Bayesian latent class analysis produced diagnostic accuracy estimates that were more interpretable than composite reference standards for extrapulmonary tuberculosis tests

5.10.1 Supplemental methods

5.10.1.1 Latent class model: Likelihood and prior distributions

Observed test results for the i^{th} subject (*i = 1,…, N*) on the j^{th} diagnostic test (*j=1,…,4*) are denoted by T_{ij}. Test results take values 1 (for a positive test) and 0 (for a negative test).

Likelihood:

We assume:

T_{ij} ~ Bernoulli(p_{ij}), where $p_{ij} = \sum_{Li=1}^{4} \Pr(L_i) \Pr(T_{ij} = 1 | L_i)$

and where L_i denotes the latent classes of interest, such that $\sum_{Li=1}^{4} \Pr(L_i) = 1$. $L_i = 1$ denotes EPTB-positive, non-specific measurand-positive, $L_i = 2$ denotes EPTB-positive, non-specific measurand-negative, $L_i = 3$ denotes EPTB-negative, non-specific measurand-positive, and $L_i = 4$ denotes EPTB-negative, non-specific measurand-negative.

The probability of a positive test result conditional on each latent class is expressed as follows:

$\Pr(T_{ij} = 1 | L_i = l) = \Phi(a_{jl} + b_{jl} r_i)$, where Φ denotes the cumulative normal distribution function, a_{jl} and b_l are unknown parameters to be estimated and r_i ~ N(0, 1) is a subject-level random effect.

The following constraints were placed on the a_{jl} and b_l parameters to reflect that the accuracy of each test was only determined by its own measurand:

$a_{j1} = a_{j2}$ and $a_{j3} = a_{j4}$ for culture, Xpert and smear

$a_{j1} = a_{j3}$ and $a_{j2} = a_{j4}$ for the non-specific tests, cytopathology/histopathology and ADA

$b_{j1} = b_{j2}$ for culture, Xpert and smear for all forms of extrapulmonary TB

$b_{j1} = b_{j2}$ for cytopathology/histopathology for TB lymphadenitis

$b_{jl} = 0$ for ADA when l=1 or l=2

$b_{jl} = 0$ for all tests when l=3 or l=4

Prior distributions:

The following non-informative prior distributions were used:

L_i ~ Dirichlet((1,1,1,1))

a_{jl} ~ Normal(0,1)

b_{jl} ~ Uniform(0,5)

To avoid label switching, we truncated the prior distributions over the a_{4l} values such that they covered values greater than 0 when l=1=3 and covered values less than 0 when l=2=4.

MCMC sampling details

Three MCMC chains were run. After an adaptation phase of 2000 iterations, followed by a burn-in phase of 5000 iterations, we retained 15,000 iterations in each chain. For all parameters of interest, we examined the trace plots, the Gelman-Rubin R-hat statistic and the posterior density plots to determine if randomly selected initial values led to the same solution and convergence was achieved. Once convergence was achieved, we extracted the posterior median and equal-tailed 95% credible intervals for the parameters of interest.

Model fitting and checking

To examine how well the model agrees with the data used to fit it, we compared the different observed and expected frequencies of different patterns of test results. To verify if our approach to modeling conditional dependence was appropriate, we plotted the observed and expected pairwise correlation residuals (1). We also considered the probability of receiving treatment for extra-pulmonary TB versus the probability of extra-pulmonary TB, as estimated by the model.

Identifiability

To determine if a model was identifiable, we first checked if the number of unknown parameters was less than the available number of degrees of freedom. All models were fit with data on four observed dichotomous tests. Therefore, the number of degrees of freedom available in each case was $2^4 - 1 = 15$. For all models, the number of unknown parameters to be estimated was 12 (3 latent class probabilities, 8 a_{jl} parameters and 1 b_{jl} parameter). Therefore, the necessary condition for identifiability that the number of degrees of freedom exceed the number parameters to be estimated was satisfied.

We additionally checked for a sufficient condition for local identifiability by examining the rank of the Jacobian of the transformation from the joint probabilities of the test results to the parameters being estimated (sensitivity, specificity, prevalence parameters) (2). For the TB meningitis and TB pleuritis models we found that the criterion of local identifiability was not met as the rank was 10 (less than 12, the number of unknown parameters).

To detect possible problems with convergence and non-identifiability we examined the results of using non-informative prior distributions for all parameters. We encountered label-switching problems with some chains reaching solutions corresponding to (prevalence, sensitivity, specificity), while other reached (1-prevalence, 1-specificity, 1-sensitivity). We also noticed that the parameters of interest (prevalence of extrapulmonary TB or accuracy parameters with respect to extrapulmonary TB) appeared to converge in fewer iterations than estimates of the prevalence of the non-specific measurand and accuracy with respect to this measurand, particularly for the non-specific test. This could be because in all our models three tests measured the target condition whereas only one test measured the non-specific measurand. To fix the problem of label switching, we used truncated prior distributions over the sensitivity and specificity of the non-specific test with respect to its measurand. We also provided randomly selected starting values closer to the desired solutions. The Rhat statistics were close to 1 for all parameters once the label-switching problem was removed, suggesting good convergence. A visual examination of the posterior density plots (Figure S2 A-C) shows that most parameters had posterior density plots with a single mode. The specificity of the non-specific test with respect to its measurand was the one parameter that seemed to have multiple values with the same posterior density (Figure S2).

5.10.1.2 rjags code for model likelihood and priors – TB lymphadenitis example

```
model

# variables
#r: random effect
#p: matrix of probabilities for patient i and test j
#se: sensitivities
#d: true disease status
#sp: specificities
#a: intercept
#b: coefficient
# indices 1, …, 4 indicate culture, Xpert, smear microscopy, and ADA or cytopathology/histopathology

 {
   ##--------------------- LIKELIHOOD ----------------------------------------------------##

for (i in 1:N) {

   for (j in 1:4) {
    y[i, j] ~ dbin(p[LC[i],i, j], 1)
    pp[i, j] <- pow( p[1,i,j], equals(LC[i],1) )*pow( p[2,i,j], equals(LC[i],2) )*pow( p[3,i,j], equals(LC[i],3) )* pow( p[4,i,j],
equals(LC[i],4) )
   }

   LC[i] ~ dcat(pLC[1:L])
   r[i] ~ dnorm(0,1)

   # LATENT CLASS 1 : Target condition positive (EPTB+) and Measurand positive
   p[1, i, 1] <- phi(a[1, 1] + b.RE[1] * r[i])
   p[1, i, 2] <- phi(a[1, 2] + b.RE[2] * r[i])
   p[1, i, 3] <- phi(a[1, 3] + b.RE[3] * r[i])
   p[1, i, 4] <- phi(a[1, 4] + b.RE[4] * r[i])

   # LATENT CLASS 2 : Target condition positive (EPTB+) and Measurand negative
   p[2, i, 1] <- phi(a[2, 1] + b.RE[1] * r[i])
   p[2, i, 2] <- phi(a[2, 2] + b.RE[2] * r[i])
   p[2, i, 3] <- phi(a[2, 3] + b.RE[3] * r[i])
   p[2, i, 4] <- phi(a[2, 4] + b.RE[4] * r[i])

   # LATENT CLASS 3 : Target condition negative (EPTB-) and Measurand positive
   p[3, i, 1] <- phi(a[3, 1])
   p[3, i, 2] <- phi(a[3, 2])
   p[3, i, 3] <- phi(a[3, 3])
   p[3, i, 4] <- phi(a[3, 4])

   # LATENT CLASS 4 : Target condition negative (EPTB-) and Measurand negative
   p[4, i, 1] <- phi(a[4, 1])
   p[4, i, 2] <- phi(a[4, 2])
   p[4, i, 3] <- phi(a[4, 3])
   p[4, i, 4] <- phi(a[4, 4])
 }
```

```
##--------------------- PRIORS ------------------------------------------------------##

for (j in 1:3) {
  a[1,j]    ~ dnorm(0,1)
  a[2,j]   <- a[1,j]
  a[3,j]    ~ dnorm(0,1)
  a[4,j]   <- a[3,j]
}
a[1,4]    ~ dnorm(0,1) T(0, )    # non-specific test sensitivity truncated above 0
a[2,4]    ~ dnorm(0,1) T(,0)     # non-specific test specificity truncated above 0
a[3,4]  <- a[1,4]
a[4,4]  <- a[2,4]

b.RE[1] ~ dunif(0, 5)
b.RE[2] <- b.RE[1]
b.RE[3] <- b.RE[1]
b.RE[4] <- b.RE[1]


pLC[1:L] ~ ddirch(prior[1:L])
for (i in 1:L) {
  prior[i]<-1
}

##--------------------- PARAMETERS OF INTEREST ---------------------------------------##

# SENSITIVITY AND SPECIFICITY WITH RESPECT TO TARGET CONDITION (EPTB)

# culture
se_cult <- ( phi(a[1, 1]/sqrt(1 + b.RE[1] * b.RE[1]))*pLC[1] + phi(a[2, 1]/sqrt(1 + b.RE[1] * b.RE[1]))*pLC[2]
  )/(pLC[1]+pLC[2])
sp_cult <- ( phi(-a[3, 1])*pLC[3] + phi(-a[4, 1])*pLC[4] )/(pLC[3]+pLC[4])

## xpert
se_xpert <-( phi(a[1, 2]/sqrt(1 + b.RE[1] * b.RE[1]))*pLC[1] + phi(a[2, 2]/sqrt(1 + b.RE[1] * b.RE[1]))*pLC[2]
  )/(pLC[1]+pLC[2])
sp_xpert <- ( phi(-a[3, 2])*pLC[3] + phi(-a[4, 2])*pLC[4] )/(pLC[3]+pLC[4])

# # smear
se_smear <- ( phi(a[1, 3]/sqrt(1 + b.RE[1] * b.RE[1]))*pLC[1] + phi(a[2, 3]/sqrt(1 + b.RE[1] * b.RE[1]))*pLC[2]
  )/(pLC[1]+pLC[2])
sp_smear <- ( phi(-a[3, 3])*pLC[3] + phi(-a[4, 3])*pLC[4] )/(pLC[3]+pLC[4])

#  # cytopathology/histopathology (CH)
se_CH <- ( phi(a[1, 4])*pLC[1] + phi(a[2, 4])*pLC[2] )/(pLC[1]+pLC[2])
sp_CH <- ( phi(-a[3, 4])*pLC[3] + phi(-a[4, 4])*pLC[4] )/(pLC[3]+pLC[4])

# SENSITIVITY AND SPECIFICITY WITH RESPECT TO NON-SPECIFIC MEASURAND

#  # cytopathology/histopathology (CH)
se_CH_measurand <- ( phi(a[1, 4])*pLC[1] + phi(a[3, 4])*pLC[3] )/(pLC[1]+pLC[3])
sp_CH_measurand <- ( phi(-a[2, 4])*pLC[2] + phi(-a[4, 4])*pLC[4] )/(pLC[2]+pLC[4])
```

```
  # PREVALENCE OF EPTB AND NON-SPECIFIC MEASURAND
  prev.EPTB <- pLC[1] + pLC[2]
  # # prevalence of non-specific marker
  prev.nonsp <- pLC[1] + pLC[3]
}
```

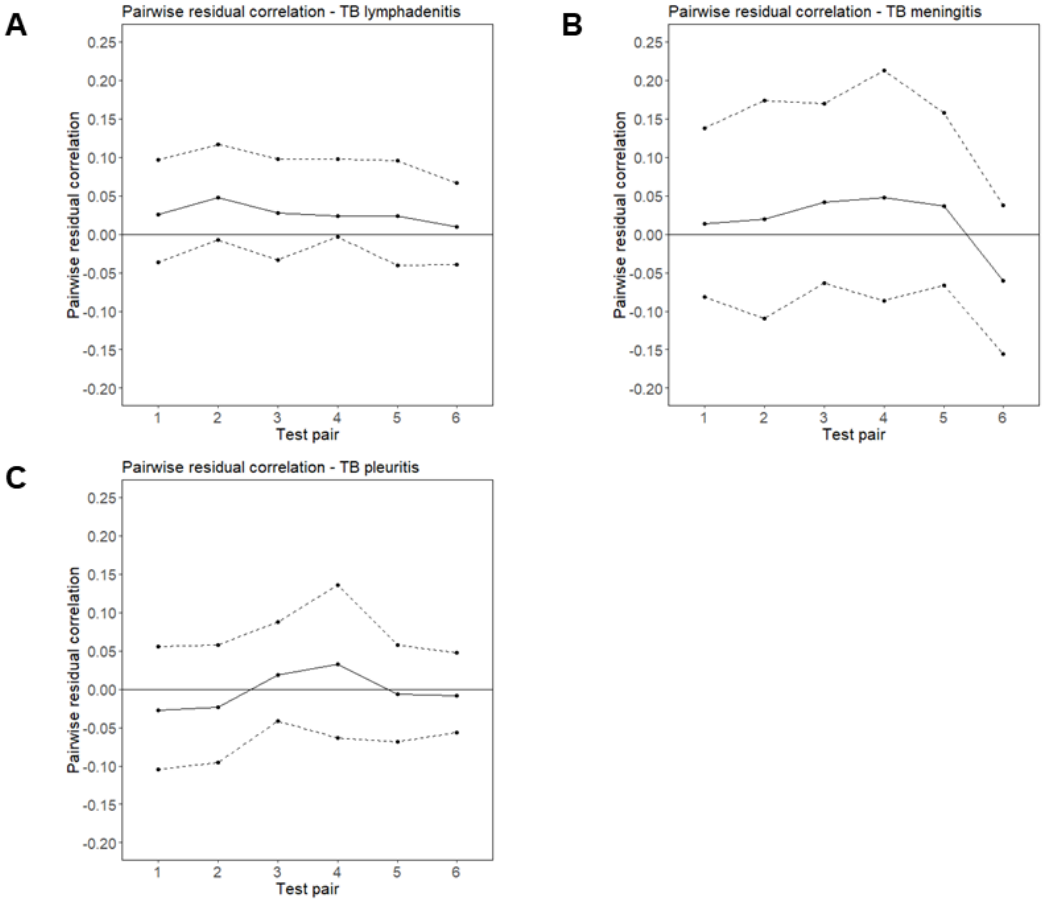
5.10.2 Supplemental results

5.10.2.1 Model fit

Table 1 displays the different test pattern results observed in lymph node samples from individuals tested for TB lymphadenitis. The observed counts of each test result pattern, along with expected counts, are shown, with generally good concordance. Exceptionally, there were fewer individuals observed with all positive test results except smear than the model expected (85 observed versus 98 expected), but the remaining patterns showed little discrepancy.

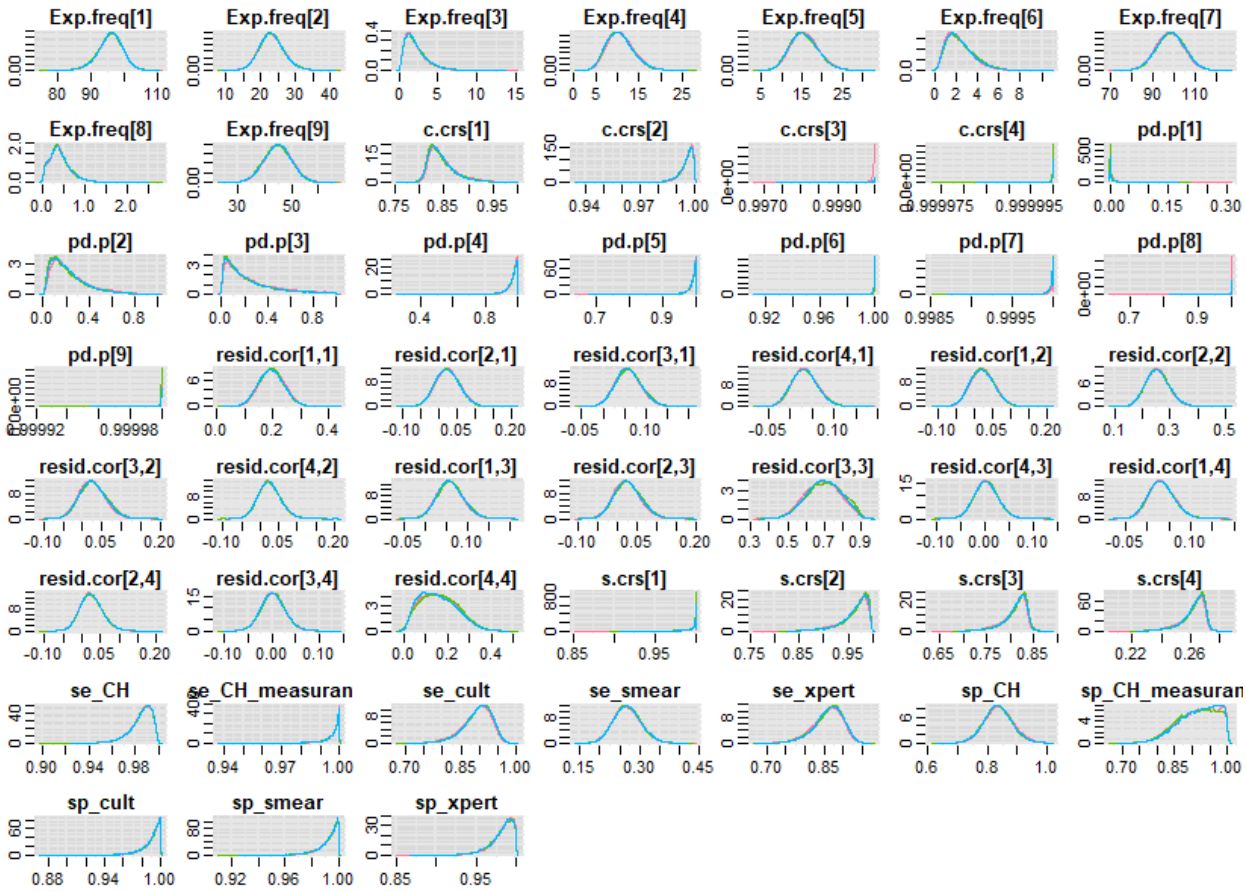
Pairwise residual correlations between test results are displayed in Supplementary figure 1. Residual correlation between tests was generally low, reflecting the observed conditional dependence was not more than expected under the model.

Figure S1: Pairwise residual correlation plots for tests run on A) lymph node, B) CSF, and C) pleural fluid. Test pair 1: culture and Xpert; 2: culture and smear microscopy; 3: culture and cytopathology/histopathology or ADA; 4: Xpert and smear microscopy; 5: Xpert and cytopathology/histopathology or ADA; 6: smear microscopy and cytopathology/histopathology or ADA. ADA – adenosine deaminase; CSF – cerebral spinal fluid; TB – tuberculosis.

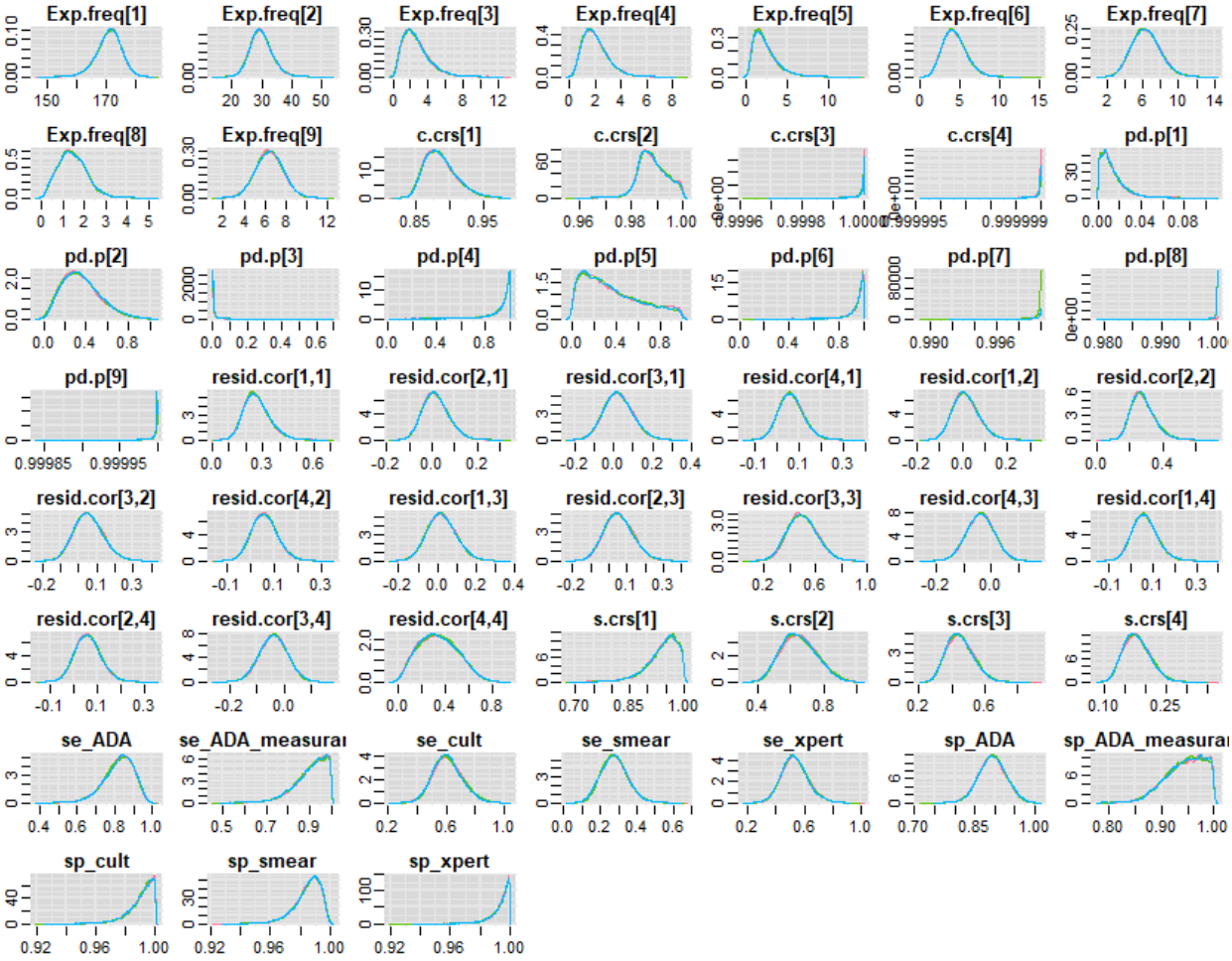


Supplemental figure S2: Density plots. For each of the three models (one per form of extrapulmonary TB), density plots were generated for each parameter of interest. For test result parameters specificity, *C*, and sensitivity, *S*, the indices 1, …,4 indicate culture, Xpert, smear microscopy, and ADA or cytopathology/histopathology.

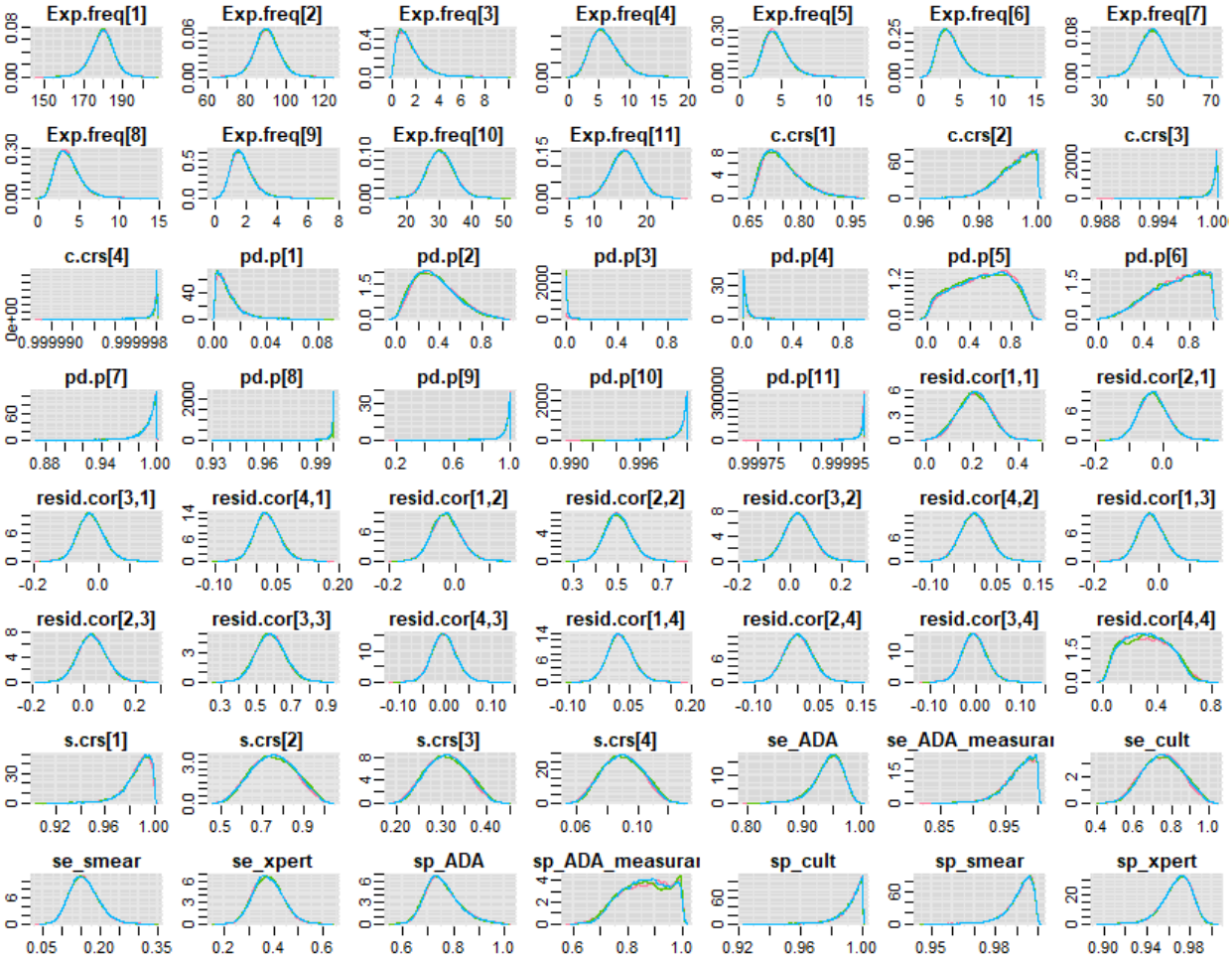
A) TB lymphadenitis



B) TB meningitis



C) TB pleuritis



5.10.2.2 Bayesian latent class analysis: estimated probabilities of the four latent classes

Supplementary Table S1: Probabilities of the four latent classes, for each form of extrapulmonary TB as estimated by Bayesian latent class analysis.

	EPTB-positive, non-specific measurand-positive ("LC1")	EPTB-positive, non-specific measurand-negative ("LC2")	EPTB-negative, non-specific measurand-positive ("LC3")	EPTB-negative, non-specific measurand-negative ("LC4")
TB lymphadenitis				
Probability (95% CrI)	0.60 (0.53, 0.67)	0.0060 (0.0004, 0.023)	0.030 (0.0010, 0.082)	0.36 (0.29, 0.43)
TB meningitis				
Probability (95% CrI)	0.14 (0.085, 0.22)	0.013 (0.001, 0.048)	0.043 (0.002, 0.13)	0.79 (0.70, 0.87)
TB pleuritis				
Probability (95% CrI)	0.31 (0.26, 0.47)	0.010 (0.001, 0.036)	0.073 (0.003, 0.21)	0.55 (0.45, 0.69)

The most common classes are EPTB-positive, non-specific measurand-positive and EPTB-negative, non-specific measurand-negative. The two classes where EPTB status and non-specific measurand are discordant were of far lower probability. This indicated that the four-class latent class model worked well to identify the target condition of interest in our study population, EPTB. CrI – credible interval; EPTB – extrapulmonary tuberculosis; TB – tuberculosis

5.10.2.3 Two-class latent class model results:

Table S2: Diagnostic accuracies of tests for each form of extrapulmonary TB by latent class analysis.

	Culture	Xpert	Smear microscopy	Cytopathology / Histopathology	ADA
TB lymphadenitis					
Sensitivity (95% CrI)	90.4 (81.9, 95.4)	86.9 (78.7, 92.4)	27.4 (21.1, 34.3)	98.9 (96.5, 99.8)	NA
Specificity (95% CrI)	99.3 (96.3,100)	98.4 (94.5,99.8)	99.4 (97.0,100)	84.8 (75.8, 95.2)	NA
TB meningitis					
Sensitivity (95% CrI)	59.2 (41.7, 80.8)	51.5 (35.0, 71.1)	27.0 (14.0, 42.2)	NA	87.8 (70.1, 97.1)
Specificity (95% CrI)	99.2 (96.8, 100)	99.5 (97.2, 100)	98.6 (96.4, 99.7)	NA	90.7 (83.8, 98.1)
TB pleuritis					
Sensitivity (95% CrI)	69.3 (51.3, 91.8)	35.0 (25.1, 47.8)	14.2 (8.63, 22.3)	NA	95.8 (90.4, 98.9)
Specificity (95% CrI)	99.4 (97.3, 100)	96.9 (93.8, 99.0)	99.3 (97.5, 99.9)	NA	79.0 (66.8, 98.6)

Table S3: Prevalence of each form of extrapulmonary TB as estimated by the two-class latent class model

	TB lymphadenitis	TB meningitis	TB pleuritis
Prevalence (95% CrI)	60.4% (54.0–67.0)	16.2% (9.65–25.0)	38.0% (27.6–51.5)

5.10.3 Supplemental references

1. Qu Y, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. Biometrics. 1996;52(3):797-810.

2. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika 1974; 61:215–231

6 CHAPTER 6: MANUSCRIPT 4: INTEGRATION OF MOLECULAR TESTING FOR *M. TUBERCULOSIS* AND SARS-COV-2 IN LIMA, PERU

6.1 Preface

The COVID-19 pandemic has led to decreased healthcare access for people with health issues unrelated to COVID-19, including TB. As described in the literature review, TB case notifications in Peru have declined precipitously since the beginning of the COVID-19 pandemic. Resources have been re-allocated towards pandemic response and away from existing health services, so it is more critical than ever to utilise the available tools and personnel in a maximally efficient manner. One approach is to integrate existing healthcare services with specific COVID-19 services. Molecular testing is one potential opportunity, particularly as Peru already has an existing network of GeneXpert platforms for TB testing.

In this manuscript, I describe our prospective study that aims to respond to the double disease burden of TB and COVID-19 by investigating the diagnostic accuracy and feasibility of integrated molecular TB and COVID-19 testing on the GeneXpert platform. This study is a collaboration between McGill University and the Universidad Peruana Cayetano Heredia.

This is a manuscript in preparation for submission to *Journal of Clinical Microbiology*.

6.2 Title page

Integration of molecular testing for *M. tuberculosis* and SARS-CoV-2 in Lima, Peru

Emily Lai Ho MacLean[a,b], Luz Villa-Castillo[c], Tatiana Caceres[c,d], Nandini Dendukuri[e], Claudia M. Denkinger[f], Giorgia Sulis[a], Mikashmi Kohli[g], Madhukar Pai[a,b], César Ugarte-Gil[c,d]

[a] Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Canada

[b] McGill TB Centre, Research Institute of the McGill University Health Centre, Montreal, Canada

[c] Instituto de Medicina Tropical Alexander von Humboldt, Universidad Peruana Cayetano Heredia Lima, Peru

[d] School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru

[e] Department of Medicine, McGill University Health Centre; Montréal, Canada.

[f] Division of Tropical Medicine, Center of Infectious Diseases, University of Heidelberg, Heidelberg, Germany

[g] FIND, Geneva, Switzerland

6.3 Abstract

Background

Tuberculosis (TB) case notification rates have plummeted since the beginning of the COVID-19 pandemic. Integrated molecular testing is an opportunity to detect and provide care for both diseases. Many high TB burden countries already have existing GeneXpert networks for TB testing using Xpert MTB/RIF Ultra (Ultra), and now a GeneXpert SARS-CoV-2 assay, Xpert Xpress SARS-CoV-2 (Xpress), is available. We evaluated the diagnostic accuracy of Xpress for SARS-CoV-2 detection and assessed the feasibility of integrating TB and COVID-19 testing using a single sputum specimen with Ultra and Xpress in Lima, Peru.

Methods

We recruited 506 consenting adults presenting with clinical symptoms or epidemiologic history suggestive of TB and/or COVID-19. All participants provided one nasopharyngeal swab (NPS) and one sputum sample. For COVID-19, NPS and sputum were tested on Xpress and standard-of-care RT-PCR; sputum was also tested for TB using culture and Ultra. Diagnostic accuracy of Xpress on sputum and NPS was estimated using different reference standards. To assess testing integration, the diagnostic yields of Xpress and Ultra were compared to that of RT-PCR on NPS and sputum culture, respectively.

Results

Data were available from 500 participants. In-study TB positivity was 9.6% (48/500) and for SARS-CoV-2 was 29% (145/500). Among TB cases, concurrent SARS-CoV-2 positivity ranged from 0% to 17%, depending on the COVID-19 test and specimen used. Compared to all reference standards, Xpress' positive percent agreement was lower on sputum than on NPS, while Xpress' negative percent agreement was higher on sputum than on NPS. Regarding integrated testing's diagnostic yield, Ultra detected 98% of culture-confirmed TB cases, while Xpress on sputum detected 83% of COVID-19 cases defined by positive RT-PCR on NPS. Using Xpress with NPS instead of sputum increased yield to 90%, a non-statistically significant difference (p=0.127). Clinical and laboratory staff reported that molecular testing for both diseases was easy and acceptable.

Conclusion

Integrated molecular testing using GeneXpert identified most cases of TB and of COVID-19 and was considered feasible, but sputum was not the ideal sample. More research is needed to determine in which situations integrated testing is most worthwhile.

6.4 Introduction

Tuberculosis (TB) was the leading cause of death by an infectious disease until the emergence of coronavirus disease 2019 (COVID-19). The World Health Organization (WHO) estimated that there were 10 million cases of TB and 1.5 million TB-related deaths in 2020 (1). Since it emerged, there have been over 485 million cases and 6 million deaths attributable to COVID-19 (2). Resource mobilisation in the face of the ongoing pandemic has been vast (3), but some of this has come at the expense of other diseases' control programs, with existing personnel, facilities, and supplies re-allocated to the COVID-19 response. Routine health services have also been disrupted due to lockdowns (4-7). For TB, this has resulted in case notification rates that were on average 23% lower in 2020-2021 compared to pre-pandemic figures (8). Since an estimated 4 million people who developed TB in 2020 were not diagnosed or notified (1) and, consequently, did not receive proper care, strategies and catch-up efforts to mitigate the detrimental effects of COVID-19 are urgently needed.

In response, in mid-2021 Stop TB Partnership and USAID recommended simultaneous, integrated (on a multiplex platform) testing approaches for TB and COVID-19 in high TB burden countries for individuals presumed to have either disease (9). In late 2021, the Global Fund released a briefing note recommending that individuals whose clinical signs and symptoms meet the case definitions for TB and COVID-19 be tested for both diseases using sputum and nasopharyngeal swabs (NPS), respectively (10). WHO has yet to issue a policy recommendation on integrated testing.

One platform that may be used for this purpose is GeneXpert (Cepheid, USA), which runs cartridge-based, automated PCR tests for a variety of diseases, including TB and COVID-19. Xpert MTB/RIF and newer generation Xpert MTB/RIF Ultra (Ultra) are WHO-endorsed molecular TB tests that run on sputum samples (11). The Xpert Xpress SARS-CoV-2 (Xpress) test has received United States Food and Drug Administration Emergency Use Authorisation for SARS-CoV-2 detection with NPS samples (12). Both Ultra and Xpress cartridges are available to National TB Programs at concessional prices via the Global Drug Facility of the Stop TB Partnership (13). Therefore, Ultra and Xpress testing on the GeneXpert platform presents a clear opportunity for presumptive patients to receive care for TB and COVID-19 in one clinical encounter, as existing equipment, staff, and know-how can be leveraged.

This type of integrated intervention could be particularly applicable in countries like Peru, a high TB burden country (14) that has been particularly hard-hit by the COVID-19 pandemic and has

one of the world's highest COVID-19 cumulative mortality rates (15). Peru has also experienced a pandemic-associated drop-off in TB case notifications: there were 8093 fewer TB case notifications from March to October 2020 than in that same period in 2019, a drop of about 20% (16). Since early 2020, there have been approximately 3.5 million notified cases of COVID-19 and 212,000 COVID-19 related deaths in Peru (2) (Figure 6.1). Due to high endemic levels of TB and multi-drug resistant TB, there is an existing network of GeneXpert platforms across seven regions of the country, and use of Xpert MTB/RIF or Ultra is recommended for TB diagnosis among key populations, such as children and people living with HIV (17).

Against this backdrop, we aimed to investigate integrated TB and COVID-19 molecular testing, i.e., simultaneously testing for both diseases in the same clinical encounter, including with a single sputum specimen, on the GeneXpert system. To do so, we estimated the accuracy of Xpress on sputum samples for COVID-19 diagnosis, since sputum is not the preferred sample for COVID-19 detection. We compared the proportion of TB and COVID-19 cases identified using Ultra and Xpress, respectively, when integrated testing is in place versus standard-of-care methods. We also comment on the feasibility of performing integrated testing.

6.5 Methods

6.5.1 Study design and population

We conducted a cross-sectional diagnostic accuracy study with prospective recruitment.

Participants were recruited from February to October 2021. Before July 2021, participants were recruited from three sites across Lima: the COVID-19 clinic at Huaycán Hospital, a secondary referral hospital in the Ate-Vitarte district; the TB clinic at Huascar XV Health Centre, a primary health facility in San Juan de Lurigancho district; and the TB clinic at Max Arias Health Centre, a primary health facility in La Victoria district. After July 2021, participants were also recruited from the TB clinics of 32 other primary health facilities in San Juan de Lurigancho district (Figure 6.1Figure 2.1).

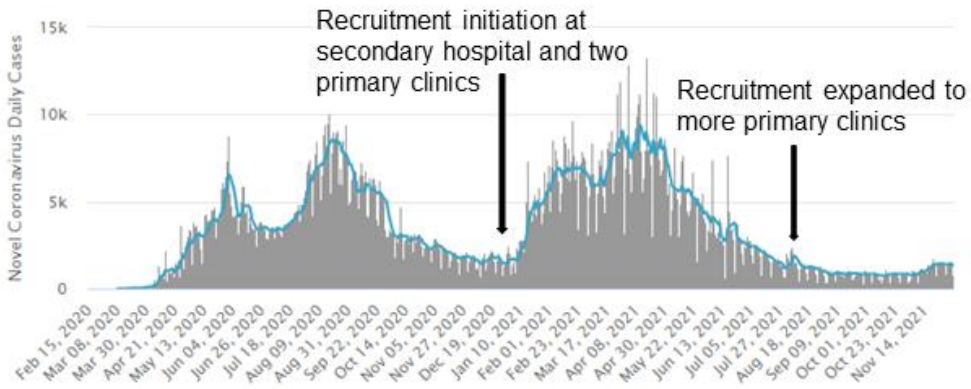


Figure 6.1: Number of daily incident COVID-19 cases in Peru between February 15 and November 14, 2021. The blue line represents weekly average. Modified from (18)

Adult outpatients (18 years or older) with presumptive COVID-19 and/or TB (i.e., symptoms including but not limited to cough, fever, difficulty breathing, sore throat of any duration; or epidemiologic history suggestive of COVID-19 and/or TB) with no history of COVID-19 in the previous three months and no history of anti-TB therapy in the previous six months were eligible for study inclusion. Participants self-reported all information.

Participants provided one spontaneously expectorated sputum sample (approximately 5 mL) as well as one NPS which was placed in approximately 3 mL of transport media. All participants provided written informed consent. This study received ethical approval from the Comité Institucional de Ética en Investigación at Universidad Peruana Cayetano Heredia (UPCH) and the McGill University Health Centre Research Ethics Board. This study was registered in the PRISA repository at Instituto Nacional de Salud in Peru (number EI00000001484).

6.5.2 Procedures

Eligible individuals presenting at a recruitment site were invited to participate in our study. Due to logistical considerations and staff availability, consecutive sampling was not feasible. NPS and sputum were stored and transported at 2-8°C and processed within the same day. Procedures are shown in Figure 6.2.

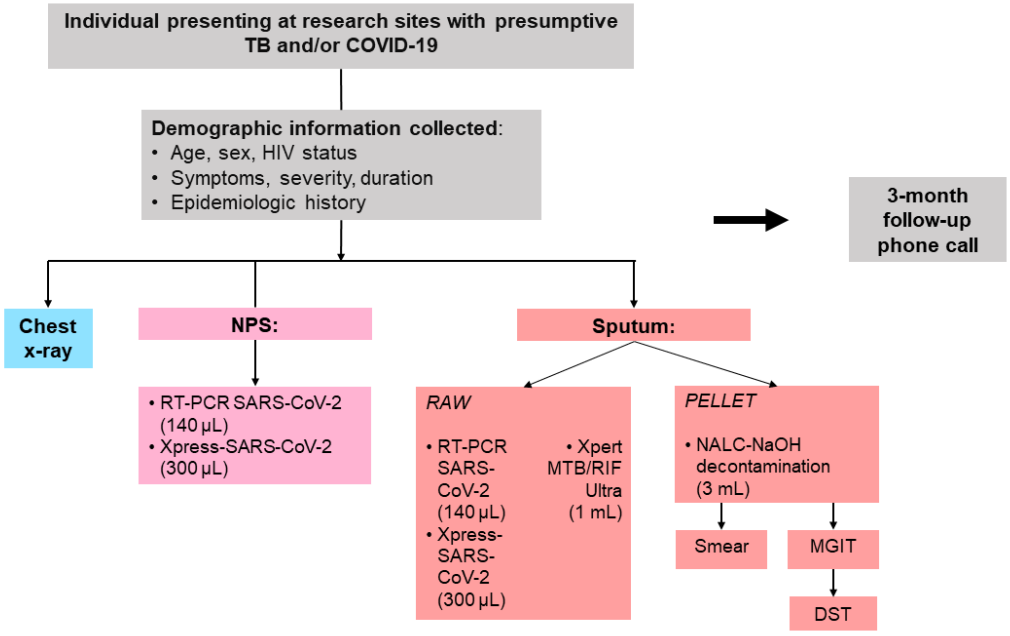


Figure 6.2: Study procedures performed for each participant recruited in the study. Each participant provided one NP swab and one sputum sample. DST – drug susceptibility testing; MGIT – Bactec MGIT culture; NPS – nasopharyngeal swab; TB – tuberculosis

NPSs were transported via approximately 3 mL of transport media (VTM Swab Kit [Boenmed, China], AB Transport Medium [AB Medical, Korea], or VTM transport medium [Liofilchem, Italy], depending on availability). 140 µL of NPS sample media was used for testing with a commercially available COVID-19 test (RT-PCR) (Norgen Biotek, Canada) according to the manufacturer's instruction for use (LOT numbers: 598068, 598069, 598323, 600755). The RT-PCR on NPS was the standard-of-care for SARS-CoV-2 detection in our study setting. A further 300 µL was used for testing on Xpress as per manufacturer's protocol (LOT numbers: 1000224306, 1000265143). RT-PCR primers target the nucleocapsid genes N1 and N2, while Xpress primers target genes for N2 and E, the envelope protein E (19, 20).

The 5 mL sputum sample was homogenized using glass beads and split. For TB testing, 1 mL was used for Ultra, while for COVID-19, 140 µL was run on RT-PCR and 300 µL was run on Xpress. The remaining sputum was decontaminated with the NALC-NaOH method; this sample was used for smear microscopy and bacteriologic culture (BACTEC MGIT, BD, USA), with positive tests undergoing drug-susceptibility testing with DST SIRE (BD).

Participants recruited at primary health clinics received chest X-rays on-site; other individuals were referred to private clinics and given a voucher to recoup costs of the procedure. Chest X-

rays were read by a blinded physician experienced in reading radiographic images. Chest X-ray abnormalities were considered suggestive of TB, COVID-19, or neither.

The laboratory staff who performed all assays were blinded to clinical details of participants. Laboratory procedures were conducted at the Dr Humberto Guerra Alisson laboratory, a reference level laboratory at the Instituto de Medicina Tropical Alexander von Humboldt at UPCH. Study data were collected and managed using REDCap electronic data capture tools hosted at the Research Institute of the McGill University Health Centre (21). REDCap (Research Electronic Data Capture) is a secure, web-based application designed to support data capture for research studies.

Information on testing integration feasibility was collected through semi-structured interviews with study staff.

6.5.3 Statistical analyses

6.5.3.1 Diagnostic accuracy of Xpress on NPS and on sputum

Contingency tables for Xpress accuracy using NPS and sputum were constructed. Positive percent agreement (PPA) and negative percent agreement (NPA) with 95% confidence intervals (95% CI) calculated by exact method were estimated compared to two reference standards for COVID-19 diagnosis: the first was defined as a positive result on RT-PCR with NPS; the second was a composite reference standard (CRS) defined as positive RT-PCR on NPS accompanied by chest radiography abnormalities suggestive of COVID-19. Tests that produced errors or inconclusive results were excluded from analyses. Analyses were conducted using `epiR` package (version 2.0.33) (22) in RStudio (23).

6.5.3.2 Proportion of TB and COVID-19 cases identified using GeneXpert

Here, the index test for TB is Ultra. There is no perfect reference standard for TB, but microbiological confirmation of *Mycobacterium tuberculosis* (*M. tuberculosis*) by liquid culture is acknowledged to be the most accurate option (24) and positivity served as reference standard for confirmed TB. The index test of interest for COVID-19 is Xpress-sputum and secondarily Xpress-NPS. As standard-of-care in our study setting, the RT-PCR on NPS test served as COVID-19 reference standard and individuals with a positive result were classified as having COVID-19. Diagnostic yield was calculated as the proportion of index test positives compared to reference standard positive. Overlap in test results was visualized using ‘VennDiagram’ (version 1.7.1) (25) in Rstudio.

6.6 Results

6.6.1 Study participant information

Between January and October, 506 eligible individuals were enrolled in our study, with 6 excluded due to inadequate sputum specimen volume. Thus, 500 participants were included in our analyses. Demographic characteristics of the study population are shown in Table 6.1. Approximately one quarter (139/500, 28%) of our study population had previously been tested for SARS-CoV-2. The study population was balanced between men and women and was relatively young (median age 39.5, range 18 to 87). Few participants (58/500, 4.7%) reported underlying comorbidities. Almost all participants were mildly ill when enrolled (457/500, 91%). The most commonly reported symptoms at study enrolment were cough (94%, 78% less than two weeks in duration, 16% longer than two weeks), headache (86%), and general malaise (83%). To be eligible for study inclusion, participants had to produce a sputum sample, thus there was likely a selection bias against the inclusion of asymptomatic individuals who had difficulty expectorating.

Table 6.1: Demographic characteristics of study population (n=500)

Trait	
Gender (%)	
Women	248 (50)
Men	252 (50)
Age (median [IQR])	39.5 (29, 54)
Known exposure to COVID-19 case in 2 weeks prior to symptom onset (%)	166 (34)
Disease severity at study entry (%)	
Not ill	0 (0)
Mildly ill	457 (91)
Moderately ill	41 (8.3)
Gravely ill	0 (0)
NA	2 (0.41)
Smoking status (%)	
Current	30 (6.2)
Former	126 (28)
Never routine	317 (65)
NA	2 (0.41)
Chest x-ray results (%)	
No abnormalities	57 (11)
Abnormal signals considered suggestive of COVID-19	162 (32)
Abnormal signals considered suggestive of TB	62 (12)

Not available	219 (44)
Comorbidities (%)	
Cardiovascular disease	29 (6.0)
Asthma	24 (5.0)
Other chronic respiratory condition	3 (0.62)
Diabetes mellitus	9 (1.9)
Known HIV infection	6 (1.2)
Chronic kidney disease	2 (0.41)
Current cancer	2 (0.41)
Symptom positivity (%)	
Breathing difficulties	197 (39)
Cough	468 (94)
Less than 2 weeks	388 (78)
2 weeks or longer	80 (16)
Fatigue	361 (72)
Fever	294 (59)
General malaise	415 (83)
Headache	429 (86)
Loss of smell	96 (19)
Loss of taste	90 (18)
Muscle pain	388 (78)
Sore throat	425 (85)
Thoracic pain	333 (67)

[a]Clinical severity definitions are as follows: Not ill: healthy and strong impression throughout examination Mildly ill: able to carry out routine activities, but symptomatic with fatigue, cough, etc. upon careful inspection. Moderately ill: some impairment of activities; visibly ill to a lay person; still ambulatory and mostly self-sufficient, but clearly symptomatic. Gravely ill: unable to carry out usual activities, visibly distressed, requires hospitalization. IQR – interquartile range; NA – not available

6.6.2 Prevalence of test-positivity

In our study population, 48/500 (9.6%, 95% CI, 7.2 to 12) of participants were found to have culture-confirmed TB, while 145/500 (29%, 95% CI, 0.25 to 0.33) were deemed positive for COVID-19 based on detection of SARS-CoV-2 using RT-PCR with NPS.

Altogether, 10 people were concurrently positive on TB culture and at least one COVID-19 test. Among the 48 people with culture-confirmed TB, the number who also tested positive for SARS-CoV-2 varied by test and specimen (Table 6.2): for RT-PCR, 3 tested positive on NPS (6.3%) and none tested positive on sputum (0.0%), while for Xpress, 8 tested positive on NPS (17%) and 2 tested positive on sputum (4.2%). This demonstrates that the rate of concurrent TB and COVID-19 observed in a particular population depends on the test and specimen used for SARS-CoV-2 detection. From the alternative perspective, of 145 people with standard-of-care-defined COVID-19, 3 individuals (2.1%) had culture-confirmed TB.

Table 6.2: Test result patterns among 48 people with culture-confirmed TB. The number of times the test result pattern was observed is indicated in the right-most column.

Test result					Observed frequency
Culture	RT-PCR on NPS for SARS-CoV-2	RT-PCR on sputum for SARS-CoV-2	Xpress on NPS for SARS-CoV-2	Xpress on sputum for SARS-CoV-2	
+	-	-	-	-	38
+	-	-	-	+	1
+	-	-	+	-	6
+	+	-	-	-	1
+	+	-	+	-	1
+	+	-	+	+	1

"+" denotes a positive test result and "-" denotes a negative test result. NPS – nasopharyngeal swab; Xpress – Xpert Xpress SARS-CoV-2

Symptoms are stratified by disease status in Table 6.3. Participants ultimately diagnosed with COVID-19 had higher rates of most symptoms, including fatigue, headache, loss of smell or taste, and muscle pain. However, individuals diagnosed with TB typically reported experiencing a particular symptom longer than those with COVID-19. For example, breathing difficulties were reported in 46% of participants with TB and in 48% of participants diagnosed with COVID-19; but the median duration among people with TB was 7 days (IQR: 2, 14) compared to 3 days (IQR: 2, 4) among people with COVID-19. Cough lasting longer than two weeks was very rare in people with COVID-19 (7/145, 5%) but was observed in almost half of people with TB (23/48, 48%) (Figure 6.3). Yet, cough for less than 2 weeks could not rule-out either disease, as 91% of people with COVID-19 and 42% of people with TB reported short-term cough. Symptoms are also presented for the 10 people with culture-confirmed TB and a positive result on any COVID-19 test. This is a small group (1% of study population) and therefore strong inferences should not be drawn.

Table 6.3: Time since symptom onset in 48 people with TB, as defined by culture-positivity, and in 145 people with COVID-19, as defined by NP swab RT-PCR positivity.

Trait	No. people with TB reporting symptom (%)	Symptom duration in people with TB (median days [IQR])	No. people with COVID-19 reporting symptom (%)	Symptom duration in people with COVID-19 (median days [IQR])	No. people with TB and any concurrent positive COVID-19 test (%)	Symptom duration in people with TB and any concurrent positive COVID-19 test (median days [IQR])
Breathing difficulties	22 (46)	7 (2, 14)	69 (48)	3 (2, 4)	5 (50)	7 (4, 7)
Cough	43 (90)	15 (8, 32)	139 (96)	5 (3, 8)	9 (90)	14 (7, 19)
Less than 2 weeks	20 (42)	8 (5, 9)	132 (91)	5 (3, 7)	4 (40)	6 (4, 8)
2 weeks or longer	23 (48)	32 (20, 46)	7 (5)	32 (23, 35)	5 (50)	19 (15, 20)
Fatigue	26 (54)	9 (5, 15)	113 (78)	4 (2, 6)	5 (50)	5 (4, 7)
Fever	34 (71)	4 (2, 8)	101 (70)	2 (2, 4)	4 (40)	3 (2, 4)
General malaise	33 (69)	8 (4, 14)	124 (86)	4 (3, 7)	7 (70)	5 (4, 7)
Headache	29 (60)	7 (4, 14)	128 (88)	4 (2, 5)	5 (50)	6 (5, 13)
Loss of smell	3 (6.3)	5 (4, 6)	50 (35)	3 (2, 5)	1 (10)	3 (3, 3)
Loss of taste	5 (10)	4 (3, 5)	41 (28)	4 (2, 5)	1 (10)	3 (3, 3)
Muscle pain	29 (60)	8 (7, 11)	116 (80)	4 (3, 6)	5 (50)	5 (4, 7)
Sore throat	39 (81)	8 (7, 26)	124 (96)	4 (3, 6)	9 (90)	7 (5, 7)
Thoracic pain	28 (58)	11 (7, 18)	90 (62)	4 (2, 5)	7 (70)	7 (7, 11)

IQR – interquartile range; No. – number of; NPS – nasopharyngeal swab; TB – tuberculosis

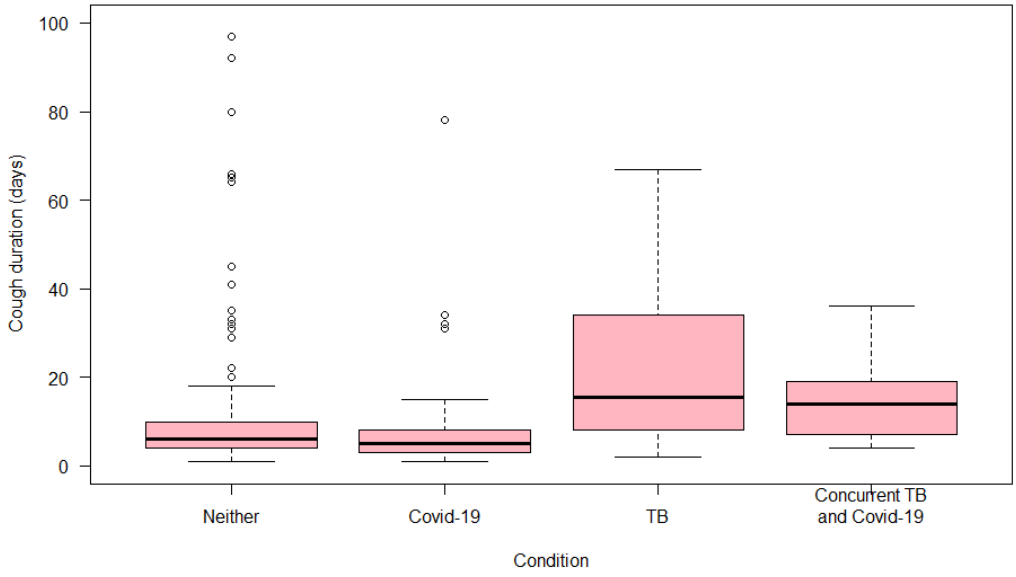


Figure 6.3: Boxplots of cough duration stratified by disease condition. COVID-19 was defined by a positive test result on RT-PCR with NP swabs (n=145). TB was defined by culture positivity (n=48). Concurrent TB and COVID-19 was defined as concurrent TB culture and a positive result on any COVID-19 test (n=10)

6.6.3 Diagnostic accuracy of Xpress on NP swab and on sputum

Estimates of Xpress's diagnostic accuracy with NPS and sputum compared to different reference standards are shown in Table 6.4. The PPA of Xpress on NPS was higher in both comparisons than Xpress on sputum. Xpress on sputum, however, had better NPA with both reference standards than Xpress on NPS. Compared to Xpress on NPS, Xpress on sputum had low PPA of 68% (95% CI, 61 to 75), although NPA was very high at 97% (95% CI, 94 to 99).

Table 6.4: Diagnostic accuracies of Xpress on NPS and Xpress on sputum for SARS-CoV-2 detection. Positive chest radiography results were any abnormal signals considered suggestive of COVID-19.

Comparison	Positive percent agreement (95% CI)	Negative percent agreement (95% CI)
Xpress (NPS) vs RT-PCR (NPS)	90 (84, 95)	82 (77, 86)
Xpress (NPS) vs RT-PCR (NPS) and chest radiography	92 (83, 97)	80 (76, 84)
Xpress (sputum) vs RT-PCR (NPS)	83 (76, 89)	93 (90, 96)
Xpress (sputum) vs RT-PCR (NPS) and chest radiography	85 (75, 92)	92 (89, 95)
Xpress (sputum) vs Xpress (NPS)	68 (61, 75)	97 (94, 99)

NPS – nasopharyngeal swab; Xpress – Xpert Xpress SARS-CoV-2

6.6.4 Diagnostic yield

All individuals included in our study were able to produce a sputum sample. Compared to culture-positive TB cases, Ultra on sputum correctly detected 46/47 (97.9%) of TB cases. Compared to individuals positive for SARS-CoV-2 on standard-of-care RT-PCR on NPS, Xpress on sputum detected 116/139 (83.4%), while Xpress on NPS detected 130/144 (90.3%), a difference that was not statistically significant (p=0.127) (Figure 6.4). Discrepant denominators are due to errors or indeterminate results: for RT-PCR on NPS, RT-PCR on sputum, Xpress on NPS, and Xpress on sputum there were 17, 14, 1, and 10 unavailable results, respectively. As 10 of our 500 participants had culture-positive TB and a positive result on any COVID-19 test, at least 100 people with presumptive TB and/or COVID-19 would have to be tested to find one person with both diseases.

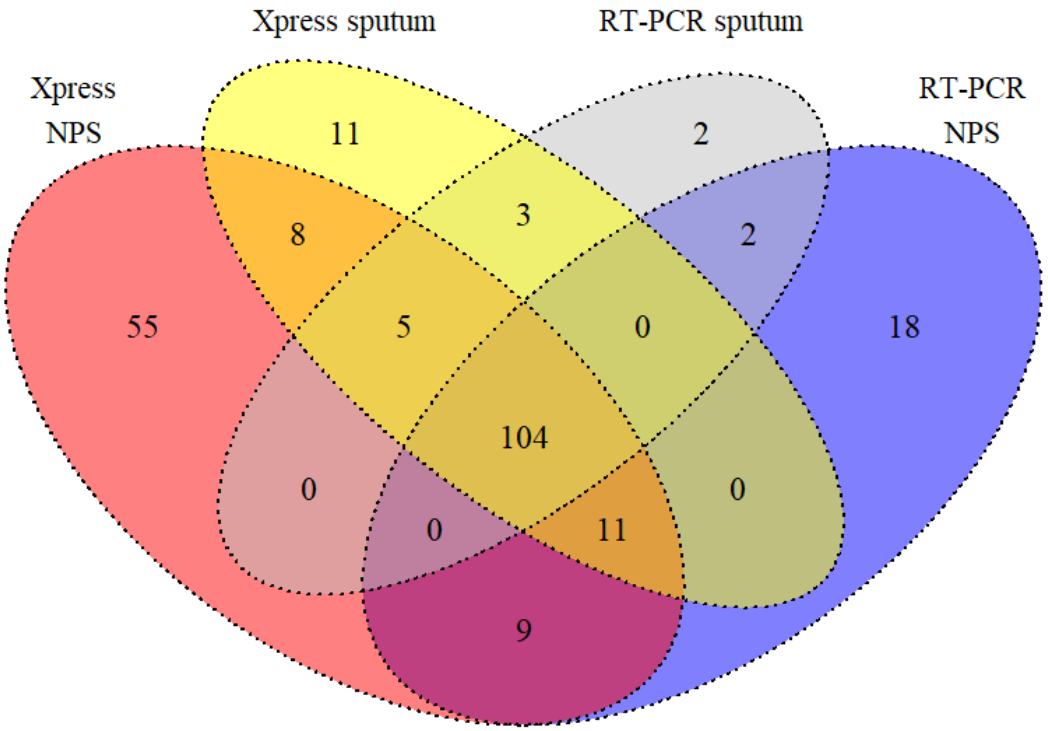


Figure 6.4: Venn diagram of COVID-19 molecular test results. NPS – nasopharyngeal swab. Xpress – Xpert Xpress SARS-CoV-2

6.6.5 Testing integration feasibility

Integration of TB and COVID-19 testing using a single sputum specimen was considered very feasible in our study. The additional collection of an NPS was not considered difficult. Clinical staff at recruitment sites were already experienced at collecting sputum samples and NPSs and

had existing protocols regarding personal protective equipment and biosafety for these procedures. Collecting NPS was considered simpler than sputum, as some participants needed coaching. Sputum specimen and NPSs were collected outdoors or in well-ventilated areas to reduce biosafety concerns. Splitting the sputum sample for use in multiple assays may be challenging, but laboratory staff used glass beads to homogenise the sample prior to allocating the appropriate volume to each test. As UPCH already utilizes Ultra in their laboratory work-flow, the addition of Xpress was straight-forward, whether using sputum or NPS.

6.7 Discussion

Even as the COVID-19 pandemic continues, it is crucial that on-going public health threats like TB are not ignored. In high TB burden settings, taking advantage of already widespread testing platforms to concurrently investigate TB and COVID-19 could be one way to start recovering the millions of people with TB who have not received care since the pandemic's onset. Finding these missing cases is especially critical given that recent evidence suggests that people with previously diagnosed or current TB are at high risk of complications and death if infected by SARS-CoV-2. A multi-country cohort investigating the effects of COVID-19 in 767 people with current or previous TB reported that 61.7% were hospitalized for COVID-19 and more than 1 in 10 died (11.1% (85/767)) (26). Recent data from Peru suggest that mortality risk in people with concurrent TB and COVID-19 is 7.3%, compared to 4.6% in people with TB alone (16).

As such, our study investigated the integration of COVID-19 and TB testing, namely by simplifying the identification of both diseases by using a single respiratory sample for testing on the GeneXpert platform. Compared to our setting's standards-of-care, this approach had a diagnostic yield of 97.9% for TB and 83.4% for COVID-19. Using NPS raised the proportion of COVID-19 cases identified to 90.3%. We found that Xpress had moderate PPA and relatively high NPA on sputum, and relatively high PPA and low NPA on NPS.

Based on our findings from 2021 in Lima, using a single sputum specimen on GeneXpert may not be the ideal approach for COVID-19 and TB testing integration, since Xpress had moderate PPA on sputum. Indeed, our PPA estimates for the molecular tests were lower than have been previously reported (27, 28). In a hypothetical population of 1000 people diagnosed with COVID-19 based on positive RT-PCR with NPS, Xpress on sputum would miss 166 individuals, while Xpress on NPS would miss 97. It should be noted that these figures are the result of a comparison to one particular definition of COVID-19, and therefore the exact numbers would

vary by reference standard deployed. Consider that Xpress produced more positive results than RT-PCR with either NPS or sputum, but our reference standard definitions deemed those additional positive results as ‘false positives’. As the reference standard is considered perfect, all observed error, i.e., discrepant results, are attributed to Xpress (29, 30). However, Xpress likely has improved analytical characteristics compared to the reference standards. Analytical studies have reported Xpress’ limit of detection as ranging from 8.26 copies/mL (19, 31) to 100 copies/mL (32), which is lower than RT-PCR’s at approximately 2 copies/µL (20, 33). This might explain Xpress’ lower NPA on NPS, as it may be able to detect SARS-CoV-2 genetic material in people near the end of their SARS-CoV-2 infection. Therefore, in situations without a perfect test, careful consideration of reference standard definitions is critical, and alternative analytic approaches, like latent class analysis, may be important to detect biased results.

New recommendations from The Global Fund recommend testing for COVID-19 and TB when “clinical signs and symptoms meet the case definitions for both diseases” (10). This seems consistent with our low concurrent positivity rate, suggesting that testing everyone with either presumptive TB or COVID-19 for both diseases may not be a worthwhile approach. However, 17% of people with culture-confirmed TB also tested positive for SARS-CoV-2 on Xpress with NPS, so integrated disease testing may be more pertinent in people with presumptive TB.

Unfortunately, clinical signs and symptom duration alone could not differentiate participants diagnosed with TB from people diagnosed with COVID-19 (Table 6.3). Our study population, although almost all mildly ill, were almost all symptomatic and the relatively high upper IQR values for symptom durations suggest that many had delayed seeking care, even as local COVID-19 prevalence was quite high. Many symptoms, including cough and breathing difficulties, were frequently reported by both groups, although some, including loss of smell or taste and headache, were more frequently reported by participants who were SARS-CoV-2 positive. Although median symptom duration in people with positive RT-PCR for SARS-CoV-2 was less than one week, duration distributions overlapped with those observed in people with TB. Contrastingly, the median symptom duration for people with TB was typically closer to one week or longer. In those (very few) individuals with culture-positive TB and a positive SARS-CoV-2 test, the moderately longer reported durations perhaps indicates that these individuals had incidental SARS-CoV-2 infections. And in those where neither *M. tuberculosis* or SARS-CoV-2 was detected, the intermediate symptom duration may be the result of no infection or of a prior infection that is now only present at undetectably low levels.

Therefore, from our study population, it appears that integrated TB and COVID-19 testing may be most warranted when individuals are presenting with symptoms lasting less than one week. In particular, for individuals who will receive steroids as part of COVID-19 treatment, it would be very important to know whether they also have TB. For those with longer-lasting symptoms, seemingly only TB testing is needed. More research is needed to understand if this trend is maintained in other high TB burden settings, and caution should be exercised when generalizing our findings, as they are only representative of one city during a period of high COVID-19 incidence (Figure 1). Investigating testing integration in another setting may result in different findings.

As resources have been re-allocated towards pandemic response and away from existing health services, it is necessary to utilize available tools in a maximally efficient manner. Sample collection supply shortages have driven the investigation into NPS alternatives for molecular testing (34). For example, using saliva for molecular testing had comparable accuracy (35) and would be less resource-intensive in terms of personnel and cost (36). We have shown that using a single sputum sample to test for both *M. tuberculosis* and SARS-CoV-2 in a single clinical encounter was feasible, particularly as clinical and laboratory staff already are experienced in obtaining and processing sputum samples and with the GeneXpert system. Anecdotally, individuals whose clinical picture was most suggestive of TB were pleased to be offered COVID-19 testing. This suggests integrated testing was acceptable in our setting. Other work in Kerala, India has shown that systematic integration of COVID-19 and TB testing is possible, particularly with provider-directed advocacy campaigns (37). In high TB burden settings like South Africa, it has been suggested that community-based screening networks deployed for COVID-19 could support TB testing and linkage to care (38).

In planning for future pandemics, and when considering the breadth of infectious diseases endemic to many high TB burden settings, integrated testing on a multi-disease platform seems an obvious intervention, although running multiple tests simultaneously requires resources. GeneXpert is one possible platform already available in many high TB burden countries (39), but other well-established systems, e.g., m2000 Realti*m*e System (Abbott, USA), or more novel options, e.g. Truelab (Molbio, India), exist for which TB tests are also WHO-endorsed. Most of these platforms can also evaluate other repository pathogens. It is likely particularly relevant for lower-resource settings, where budget constraints may preclude procurement of multiple stand-alone systems, to invest in platforms that can be quickly updated to incorporate assays for emerging pathogens. Diagnostic companies must ensure these tools are accessible.

6.7.1 Strength and limitations

Our study is one of the first to investigate the diagnostic accuracy of Xpress on sputum sample and its integration with TB testing in a real-life pandemic setting. Our study had a large sample size, which improves measures of uncertainty and reduces random error. We ran multiple tests for each participant. Findings on testing integration feasibility are likely generalizable to other urban settings with a high burden of TB and COVID-19.

Regarding challenges, we faced some logistical issues related to the COVID-19 pandemic. As participants needed to provide a sputum sample, our results are likely most generalizable to settings testing mostly symptomatic individuals. As well, the index text, Xpress, likely had better analytical sensitivity than the reference standard, which makes interpreting PPA and NPA challenging.

6.7.2 Outstanding issues

Further work is needed to determine whether there is some combination of clinical signs and epidemiologic history that could indicate in whom integrated TB and COVID-19 testing is most warranted, which specific contexts should prioritise integrated testing (e.g., high TB prevalence settings, populations with a high prevalence of immune-compromised individuals), and the impact integrated testing policies have on recovering missing TB patients.

6.7.3 Conclusions

The diagnostic yield of Xpress on sputum was lower than on NPS (yet non-statistically significant), but integrated testing for TB and COVID-19 using GeneXpert was feasible. However, systematic testing for both diseases may not be a worthwhile approach in people presenting with either presumptive TB or COVID-19, as concurrently positive cases were rare in our study population. More research is needed to understand when integrated testing is most appropriate and to identify the epidemiological situations wherein integrated testing would be most valuable.

FUNDING

The authors are grateful to the Canadian Institutes of Health Research and International Development Research Centre for operational funding. The authors are grateful to the Global Drug Facility for providing Xpress SARS-CoV-2 cartridges.

6.8 References

1. World Health Organization. Global Tuberculosis Report 2021. Geneva: World Health Organization; 2021.
2. Johns Hopkins University. COVID-19 dashboard by the center for systems science and engineering at Johns Hopkins University. In: University JH, editor. Coronavirus Research Centre. Baltimore, USA: Johns Hopkins University; 2021.
3. Global Preparedness Monitoring Board. A World in Disorder: GPMB 2020 Annual Report. Geneva: World Health Organization; 2020.
4. Klinton JS, Heitkamp P, Rashid A, Faleye BO, Win Htat H, Hussain H, et al. One year of COVID-19 and its impact on private provider engagement for TB: A rapid assessment of intermediary NGOs in seven high TB burden countries. J Clin Tuberc Other Mycobact Dis. 2021;25:100277-.
5. Stop TB Partnership. The Impact of COVID-19 on the TB Epidemic: A Community Perspective. Geneva: Stop TB Partnership; 2020.
6. Maurer FP, Shubladze N, Kalmambetova G, Felker I, Kuchukhidze G, Drobniewski F, et al. Impact of the COVID-19 pandemic on tuberculosis national reference laboratory services in the WHO European Region, March to November 2020. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2021;26(24):2100426.
7. Migliori GB, Thong PM, Akkerman O, Alffenaar JW, Álvarez-Navascués F, Assao-Neino MM, et al. Worldwide Effects of Coronavirus Disease Pandemic on Tuberculosis Services, January-April 2020. Emerging infectious diseases. 2020;26(11):2709-12.
8. Stop TB Partnership. 12 Months of COVID-19 Eliminated 12 Years of Progress in the Global Fight Against Tuberculosis. Geneva: Stop TB Partnership; 2021 18 March 2021. Contract No.: 2.
9. USAID, Stop TB Partnership. Simultaneous, integrated diagnostic testing approach to detect COVID-19 and TB in high TB burden countries. Geneva: Stop TB Partnership; 2021.
10. The Global Fund. Briefing Note: Testing for both Tuberculosis and SARS-CoV-2. Geneva: The Global Fund; 2021 13 October 2021.
11. World Health Organization. Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update. Geneva: World Health Organization; 2020.
12. Cepheid Receives Emergency Use Authorization from FDA for Rapid SARS-CoV-2 Test [press release]. Sunnyvale, USA: Cepheid, 21 March 2020 2020.
13. Stop TB Partnership. Global Drug Facility: Diagnostics Catalogue. Geneva: Stop TB Partnership; 2021.
14. World Health Organization. WHO global lists of high burden countries for tuberculosis (TB), TB/HIV and multidrug/rifampicin-resistant TB (MDR/RR-TB), 2021–2025. Geneva: World Health Organization; 2021.
15. Pašović M, Leach-Kemon K, Troeger C, Vos T, Lozano R. Countries Hit Hardest by COVID-19 Washinton DC: Think Global Health; 2021 [cited 2021 23 November]. Available from: https://www.thinkglobalhealth.org/article/countries-hit-hardest-covid-19.
16. Ugarte-Gil C, Curisinche M, Figueroa CJ, Gotuzzo E, Rios R. COVID-19 among tuberculosis patients in Peru: An operational report from the national registry. 52nd Union World Conference on Lung Health; 20 October 2021; Virtual: The Union for Tuberculosis and Lung Health; 2021.
17. Norma téchnica de salud para la prevencion y control de la coninfeccion tuberculosis y virus de la immunodeficiencia humane en el Peru, Stat. 143-2018 (2 October 2018, 2018).
18. Worldometers.info. Worldometer: Peru dashboard Dover, Delware, United States2021 [updated 6 December 2021; cited 2021 6 December]. Dashboard]. Available from: https://www.worldometers.info/coronavirus/country/peru/.

19. Cepheid. Instructions for use: Xpert® Xpress SARS-CoV-2. In: Inc. C, editor. Sunnyvale, USA: Cepheid Inc.; 2020.
20. Pearson JD, Trcka D, Lu S, Hyduk SJ, Jen M, Aynaud MM, et al. Comparison of SARS-CoV-2 indirect and direct RT-qPCR detection methods. Virology journal. 2021;18(1):99.
21. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics. 2009;42(2):377-81.
22. Stevenson M. Tools for the Analysis of Epidemiological Data. In: Stevenson M, editor. 2.0.33 ed. Melbourne: The University of Melbourne; 2017.
23. RStudio Team. RStudio: Integrated Development for R. In: Rstudio I, editor. Boston: RStudio, Inc.; 2016.
24. World Health Organization. Policy framework: Implementing tuberculosis diagnostics. Geneva: World Health Organization; 2015.
25. Chen H. VennDiagram: Generate High-Resolution Venn and Euler Plots. In: Boutros P, editor. 1.7.1 ed. Los Angeles: University of California Los Angeles; 2021. p. R package.
26. Group TC-GS. Tuberculosis and COVID-19 co-infection: description of the global cohort. The European respiratory journal. 2021.
27. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. The Cochrane database of systematic reviews. 2021;3(3):Cd013705.
28. Lee J, Song JU. Diagnostic accuracy of the Cepheid Xpert Xpress and the Abbott ID NOW assay for rapid detection of SARS-CoV-2: A systematic review and meta-analysis. J Med Virol. 2021;93(7):4523-31.
29. Dramé M, Tabue Teguo M, Proye E, Hequet F, Hentzien M, Kanagaratnam L, et al. Should RT-PCR be considered a gold standard in the diagnosis of COVID-19? Journal of Medical Virology. 2020;92(11):2312-3.
30. United States Food and Drug Administration. Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests - Guidance for Industry and FDA Staff. In: Services USDoHH, editor. Silver Spring, Maryland: Center for Devices and Radiological Health; 2007.
31. Wolters F, van de Bovenkamp J, van den Bosch B, van den Brink S, Broeders M, Chung NH, et al. Multi-center evaluation of cepheid xpert® xpress SARS-CoV-2 point-of-care test during the SARS-CoV-2 pandemic. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology. 2020;128:104426-.
32. Zhen W, Smith E, Manji R, Schron D, Berry GJ. Clinical Evaluation of Three Sample-to-Answer Platforms for Detection of SARS-CoV-2. J Clin Microbiol. 2020;58(8).
33. ElJilani MM, Abdusalam M, Abdalla A, Dalyoum T, Ahmad S, Alhudiri I, et al. Performance Evaluation of Seven Commercial Real-Time Polymerase Chain Reaction Assays for SARS-COV-2 Detection. Libyan J Med Sci. 2021;5(2):56-60.
34. Lee RA, Herigon JC, Benedetti A, Pollock NR, Denkinger CM. Performance of Saliva, Oropharyngeal Swabs, and Nasal Swabs for SARS-CoV-2 Molecular Detection: a Systematic Review and Meta-analysis. J Clin Microbiol. 2021;59(5).
35. Butler-Laporte G, Lawandi A, Schiller I, Yao M, Dendukuri N, McDonald EG, et al. Comparison of Saliva and Nasopharyngeal Swab Nucleic Acid Amplification Testing for Detection of SARS-CoV-2: A Systematic Review and Meta-analysis. JAMA Internal Medicine. 2021;181(3):353-60.
36. Campbell JR, Uppal A, Oxlade O, Fregonese F, Bastos ML, Lan Z, et al. Active testing of groups at increased risk of acquiring SARS-CoV-2 in Canada: costs and human resource needs. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2020;192(40):E1146-E55.

37. Khobragade RN, Murthy N, Aloysius S, Surendran D, Rakesh PS, Balakrishnan S. Experience of integrated screening and testing for TB and COVID19 from Kerala, India. Public Health Pract (Oxf). 2021;2:100198-.
38. Loveday M, Cox H, Evans D, Furin J, Ndjeka N, Osman M, et al. Opportunities from a new disease for an old threat: Extending COVID-19 efforts to address tuberculosis in South Africa. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2020;110(12):1160-7.
39. MSF. Step up for TB 2020: Tuberculosis Policies in 37 countries. Geneva: Medicins Sans Frontieres; 2020 November 2020.

7 CHAPTER 7: SUMMARY & CONCLUSIONS

7.1 Summary of results

In my review of molecular TB diagnostics, I discussed recent advances in molecular tests for the detection of TB as well as the identification of anti-TB drug resistance. WHO-endorsed molecular tests have improved speed and accuracy compared to smear microscopy and culture, and many emerging products are in the developmental pipeline. Promising trends include the increased diagnostic test development in countries with a high TB burden, such as India and China, and an increase in countries investing in WGS.

The results of my systematic review and meta-analysis demonstrated that Xpert testing using stool samples for pediatric TB had similar accuracy to Xpert on sputum compared to a microbiological reference standard. With a pooled sensitivity of 67% (95% CI, 52 to 79%) and specificity of 99% (95% CI, 98 to 99%), stool Xpert is more suited for use as a rule-in test, as it will still produce many false negative results. Sensitivity was higher among children with HIV (79% [95% CI, 68 to 87%]) than in HIV-uninfected children (60% [95% CI, 44 to 74%]). This review also identified the need for improved reporting of results with respect to age stratification.

My analysis of the performance of extrapulmonary TB diagnostic tests demonstrated the utility of Bayesian LCA in assessing diagnostic test accuracy. Results from the LCA were compared to those produced by a series of CRSs. The LCA approach considered each test's imperfect performance and between-test conditional dependence to model all tests' sensitivities and specificities without assuming any test was perfect. All CRSs in the series ignored these details and the CRS-derived results were more difficult to interpret, as it was not clear which CRS produced the correct accuracy estimate.

Finally, in our study of integrating TB and COVID-19 testing, we showed that testing for TB and COVID-19 on the GeneXpert platform using a single sputum sample was a feasible intervention. Diagnostic yield for COVID-19 was moderately high and yield for TB was very high compared to standard-of-care testing. However, we found that the diagnostic accuracy of Xpress using sputum was lower than Xpress using NP swabs. More research is needed in high TB burden settings to determine whether testing all individuals with presumptive TB or presumptive COVID-19 is an optimal approach.

7.2 Strengths and limitations

An overall strength of my thesis is the diverse nature of the work, including the deployment of different methods for estimating diagnostic accuracy. Meta-analysis via bivariate random effect models, Bayesian latent class analysis, and composite reference standards are all used in appropriate instances. All manuscript and analyses were planned with input from experts in epidemiology, biostatistics, clinical medicine, and global health. As such, the manuscripts all have broad perspective.

Of course, it is also important to acknowledge limitations. In the meta-analysis, the pooled sensitivity and specificity were generated by comparison with two reference standards that are known to be imperfect; therefore, the resultant accuracy estimates are also biased. However, as there is no perfect reference standard for pediatric TB and these two reference standards are widely understood, we accepted this degree of bias. Thus, we produced an imperfect estimate, but one that incorporated all the data available at the time. Regarding the latent class analyses, the models are not as immediately intuitive to understand as composite reference standards, but the theory behind them is well-characterised and sound. As with any modeling exercise, validation can be a challenge, but I have attempted to be transparent and clear with each model's assumptions. I also have explicitly illustrated the relationships between variables with heuristic diagrams. Additionally, generalizability of my findings is not universal; this is a direct result of my thesis' exploration of the use of molecular TB tests in very specific sub-populations. For example, the sensitivity and specificity estimates of tests used for extrapulmonary TB diagnosis are probably most applicable to tertiary hospital settings, while the findings regarding the integration of TB and COVID-19 testing with GeneXpert are most likely representative of urban settings where sample transportation is not a prohibitive issue.

7.3 Implications

In this work, I have presented evidence that supports the continued and broadened use of molecular TB diagnostics. To create inclusive policies regarding molecular testing for TB, it is necessary to understand how well tests perform in detecting TB in all its manifestations. Alternate types of TB are more difficult to diagnose accurately than PTB, and the results from my studies suggest that improved tools are needed. However, alternative analytic approaches should also be considered going forward when trying to assess the accuracy of tests for diseases without a perfect reference standard, and the work presented using LCA is one such option. My work has also shown that molecular testing for TB and COVID-19 can be integrated with minimal difficulty and that multi-disease testing on a multiplex molecular platform is feasible.

As discussed above, high confidence in the evidence base vis-à-vis diagnostic accuracy is needed for country policies to adapt and recommend their use, thus encouraging the scale-up of newer tests. As molecular tests continue to emerge from the developmental pipeline, it will be critical that they are evaluated in all people with TB, not just adults with pulmonary forms, and it will also be relevant to consider how these tests can be integrated with testing for other infectious diseases. This will help ensure policies regarding new molecular TB tests can be extended to all types of TB.

7.4 Directions for future research

Evaluating new molecular TB tests in all people with presumptive TB using appropriate statistical methodology should be a priority for researchers in the TB diagnostics space. It is widely understood that there is no perfect TB test, so methods that can account for this nuance, like latent class analysis, should be considered for future diagnostic test accuracy estimation. It would also be wise to look beyond accuracy and consider the feasibility of integrating testing with other diseases as well as clinical utility and impact (59). Future research on upcoming molecular TB tests should examine accuracy and feasibility in a broad range of use cases early in the clinical evaluation phase.

More research is needed to understand optimal integrated testing for TB and COVID-19. Concurrent positivity for *M. tuberculosis* and SARS-CoV-2 was low in our study populations, which suggested that integrated testing for TB and COVID-19 for any individual with signs or symptoms of either disease may not be the best approach. Further examination in larger populations of combinations of epidemiologic history, clinical symptoms, and distributions of symptom duration may help in drawing firmer conclusions regarding who exactly needs to be tested for both diseases. As well, examining integrated testing in a diversity of settings is needed to understand where this intervention should immediately be applied, for example, urban settings with very high TB prevalence where there are many shared risk factors for COVID-19; rural settings in high TB burden countries where integrated disease testing could greatly improve quality of care; countries with a higher proportion of people who are immunocompromised; or in times of COVID-19 surges in areas with high TB prevalence. Additionally, as new SARS-CoV-2 variants of concern arise, integrated testing policies will need to be re-examined and updated as needed. Future studies will also help to understand the effect that integrated testing policies have on recovering missing TB cases and linking people with TB to care. Careful design of studies and outcomes will need to be considered to ensure the real impact of integrated testing can be identified and measured (92-94).

Regarding work in this thesis, research continues with the integrated testing cohort in Lima. All participants are undergoing a 3-month follow-up phone call where variables regarding their health and disease status are being collected. An expansion of this cohort is planned with the intention of examining rates of TB and COVID-19 in people who have received a diagnosis of either disease. Molecular tests will be employed to make these diagnoses.

7.5 Conclusion

To increase the uptake of molecular diagnostic tests for all people with TB and in the context of other concurrent respiratory diseases, the accuracy of these tests needs to be well-characterised. The sooner test accuracy is accurately and reliable estimated, the sooner tests can be implemented and contribute to improving patient outcomes.

In this thesis, I have demonstrated different methodological approaches for such characterization. The work in this thesis is a small step towards realizing the goal of all people with presumptive TB receiving accurate, rapid molecular testing, thereby enabling them to quickly receive appropriate care and halting disease transmission. Even in the era of a global pandemic, testing and care for other diseases must continue, and new strategies such as integrated molecular testing may have a role to play.

Thanks to the COVID-19 pandemic, the critical importance of diagnostics to health is more appreciated than ever before. Countries have invested in multiplex platforms and rapid technological developments have resulted in such ‘wish-list’ assays as take-home, disposable molecular tests (95). As a field, we must push to see these innovations applied to the next-generation of patient-centred TB tests, and ensure they are quickly evaluated for use in all people with TB.

8 BACK MATTER

8.1 References

1. Stop TB Partnership. 12 Months of COVID-19 Eliminated 12 Years of Progress in the Global Fight Against Tuberculosis. Geneva: Stop TB Partnership; 2021 18 March 2021. Contract No.: 2.
2. World Health Organization. Global Tuberculosis Report 2021. Geneva: World Health Organization; 2021.
3. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. PLoS One. 2011;6(4):e17601.
4. Migliori GB, Nardell E, Yedilbayev A, D'Ambrosio L, Centis R, Tadolini M, et al. Reducing tuberculosis transmission: a consensus document from the World Health Organization Regional Office for Europe. The European respiratory journal. 2019;53(6).
5. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature genetics. 2013;45(10):1176-82.
6. Graham SM, Cuevas LE, Jean-Philippe P, Browning R, Casenghi M, Detjen AK, et al. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2015;61Suppl 3:S179-87.
7. Lewinsohn DM, Leonard MK, LoBue PA, Cohn DL, Daley CL, Desmond E, et al. Official American Thoracic Society/Infectious Diseases Society of America/Centers for Disease Control and Prevention Clinical Practice Guidelines: Diagnosis of Tuberculosis in Adults and Children. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2017;64(2):e1-e33.
8. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 3: Rapid diagnostics for tuberculosis detection. Geneva: World Health Organization; 2020.
9. MSF. Step up for TB 2020: Tuberculosis Policies in 37 countries. Geneva: Medicins Sans Frontieres; 2020 November 2020.
10. Riley RL, Mills CC, O'Grady F, Sultan LU, Wittstadt F, Shivpuri DN. Infectiousness of air from a tuberculosis ward. Ultraviolet irradiation of infected air: comparative infectiousness of different patients. The American review of respiratory disease. 1962;85:511-25.
11. Yates TA, Khan PY, Knight GM, Taylor JG, McHugh TD, Lipman M, et al. The transmission of Mycobacterium tuberculosis in high burden settings. Lancet Infect Dis. 2016;16(2):227-38.
12. Mathur P, Sacks L, Auten G, Sall R, Levy C, Gordin F. Delayed diagnosis of pulmonary tuberculosis in city hospitals. Archives of internal medicine. 1994;154(3):306-10.
13. Kantor HS, Poblete R, Pusateri SL. Nosocomial transmission of tuberculosis from unsuspected disease. The American journal of medicine. 1988;84(5):833-8.
14. Otero L, Shah L, Verdonck K, Battaglioli T, Brewer T, Gotuzzo E, et al. A prospective longitudinal study of tuberculosis among household contacts of smear-positive tuberculosis cases in Lima, Peru. BMC infectious diseases. 2016;16:259-.
15. Melsew YA, Doan TN, Gambhir M, Cheng AC, McBryde E, Trauer JM. Risk factors for infectiousness of patients with tuberculosis: a systematic review and meta-analysis. Epidemiol Infect. 2018;146(3):345-53.
16. Rouillon A, Perdrizet S, Parrot R. Transmission of tubercle bacilli: The effects of chemotherapy. Tubercle. 1976;57(4):275-99.

17. Dharmadhikari AS, Mphahlele M, Venter K, Stoltz A, Mathebula R, Masotla T, et al. Rapid impact of effective treatment on transmission of multidrug-resistant tuberculosis. Int J Tuberc Lung Dis. 2014;18(9):1019-25.
18. Drain PK, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, et al. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. Clin Microbiol Rev. 2018;31(4).
19. Kendall EA, Shrestha S, Dowdy DW. The Epidemiological Importance of Subclinical Tuberculosis. A Critical Reappraisal. Am J Respir Crit Care Med. 2021;203(2):168-74.
20. World Health Organization. WHO End TB Strategy. Geneva: World Health Organization; 2015.
21. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nature reviews Disease primers. 2016;2:16076.
22. World Health Organization. WHO global lists of high burden countries for tuberculosis (TB), TB/HIV and multidrug/rifampicin-resistant TB (MDR/RR-TB), 2021–2025. Geneva: World Health Organization; 2021.
23. World Health Organization. Global Tuberculosis Report 2020. Geneva: World Health Organization; 2020.
24. Subbaraman R, Nathavitharana RR, Mayer KH, Satyanarayana S, Chadha VK, Arinaminpathy N, et al. Constructing care cascades for active tuberculosis: A strategy for program monitoring and identifying gaps in quality of care. PLoS Med. 2019;16(2):e1002754.
25. Subbaraman R, Nathavitharana RR, Satyanarayana S, Pai M, Thomas BE, Chadha VK, et al. The Tuberculosis Cascade of Care in India's Public Sector: A Systematic Review and Meta-analysis. PLoS Med. 2016;13(10):e1002149.
26. Tuberculosis profile: India [Internet]. World Health Organization. 2021 [cited 28 November 2021]. Available from: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&entity_type=%22country%22&lan=%22EN%22&iso2=%22IN%22.
27. Dabas H, Deo S, Sabharwal M, Pal A, Salim S, Nair L, et al. Initiative for Promoting Affordable and Quality Tuberculosis Tests (IPAQT): a market-shaping intervention in India. BMJ global health. 2019;4(6):e001539.
28. Das J, Kwan A, Daniels B, Satyanarayana S, Subbaraman R, Bergkvist S, et al. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. Lancet Infect Dis. 2015;15(11):1305-13.
29. Kwan A, Daniels B, Saria V, Satyanarayana S, Subbaraman R, McDowell A, et al. Variations in the quality of tuberculosis care in urban India: A cross-sectional, standardized patient study in two cities. PLoS medicine. 2018;15(9):e1002653-e.
30. Tuberculosis profile: Peru [Internet]. World Health Organization. 2021 [cited 28 November 2021]. Available from: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&entity_type=%22country%22&lan=%22EN%22&iso2=%22PE%22.
31. Ministerio de Salud - el Perú. Avances en la Investigación y Diagnóstico de la Tuberculosis en el Perú. Lima: Ministerio de Salud - el Perú; 2018.
32. McQuaid CF, Vassall A, Cohen T, Fiekert K, White RG. The impact of COVID-19 on TB: a review of the data. Int J Tuberc Lung Dis. 2021;25(6):436-46.
33. The TB/COVID-19 Global Study Group. Tuberculosis and COVID-19 co-infection: description of the global cohort. The European respiratory journal. 2021;in press.
34. Kumar MS, Surendran D, Manu MS, Rakesh PS, Balakrishnan S. Mortality due to TB-COVID-19 coinfection in India. Int J Tuberc Lung Dis. 2021;25(3):250-1.
35. Loyola S, Valle A, Montero S, Carrasco-Escobar G. Recomendaciones para describir de forma adecuada una curva epidemica de COVID-19. Rev Peru Med Exp Salud Publica. 2020;37(2):1-3.

36. Johns Hopkins University. COVID-19 dashboard by the center for systems science and engineering at Johns Hopkins University. In: University JH, editor. Coronavirus Research Centre. Baltimore, USA: Johns Hopkins University; 2021.
37. Ugarte-Gil C. TB Case Detection Gap 2020-2019 in Peru. In: MacLean E, editor. Figure ed. Lima: Ugarte-Gil, C.; 2021. p. 1.
38. Khobragade RN, Murthy N, Aloysius S, Surendran D, Rakesh PS, Balakrishnan S. Experience of integrated screening and testing for TB and COVID19 from Kerala, India. Public Health Pract (Oxf). 2021;2:100198-.
39. Togun T, Kampmann B, Stoker NG, Lipman M. Anticipating the impact of the COVID-19 pandemic on TB patients and TB control programmes. Annals of clinical microbiology and antimicrobials. 2020;19(1):21.
40. USAID, Stop TB Partnership. Simultaneous, integrated diagnostic testing approach to detect COVID-19 and TB in high TB burden countries. Geneva: Stop TB Partnership; 2021.
41. The Global Fund. Briefing Note: Testing for both Tuberculosis and SARS-CoV-2. Geneva: The Global Fund; 2021 13 October 2021.
42. MacLean EL, Villa-Castillo L, Ruhwald M, Ugarte-Gil C, Pai M. Integrated testing for TB and COVID-19. Med (New York, NY). 2022;3(3):162-6.
43. Ruhwald M, Hannay E, Sarin S, Kao K, Sen R, Chadha S. Considerations for simultaneous testing of COVID-19 and tuberculosis in high-burden countries. The Lancet Global Health. 2022;10(4):e465-6.
44. Storla DG, Yimer S, Bjune GA. A systematic review of delay in the diagnosis and treatment of tuberculosis. BMC Public Health. 2008;8:15.
45. Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis. 2006;6(9):570-81.
46. 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.
47. Urbanczik R, Rieder HL. Scaling up tuberculosis culture services: a precautionary note. Int J Tuberc Lung Dis. 2009;13(7):799-800.
48. World Health Organization. WHO consolidated guidelines on tuberculosis: Module 4: Treatment: drug-resistant tuberculosis treatment. Geneva: World Health Organization; 2020.
49. Lessells RJ, Cooke GS, McGrath N, Nicol MP, Newell ML, Godfrey-Faussett P. Impact of Point-of-Care Xpert MTB/RIF on Tuberculosis Treatment Initiation. A Cluster-randomized Trial. Am J Respir Crit Care Med. 2017;196(7):901-10.
50. Nair D, Navneethapandian PD, Tripathy JP, Harries AD, Klinton JS, Watson B, et al. Impact of rapid molecular diagnostic tests on time to treatment initiation and outcomes in patients with multidrug-resistant tuberculosis, Tamil Nadu, India. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2016;110(9):534-41.
51. Oxlade O, Sugarman J, Alvarez GG, Pai M, Schwartzman K. Xpert(R)MTB/RIF for the Diagnosis of Tuberculosis in a Remote Arctic Setting: Impact on Cost and Time to Treatment Initiation. PLoS One. 2016;11(3):e0150119.
52. Boyd R, Ford N, Padgen P, Cox H. Time to treatment for rifampicin-resistant tuberculosis: systematic review and meta-analysis. Int J Tuberc Lung Dis. 2017;21(11):1173-80.
53. Walzl G, McNerney R, du Plessis N, Bates M, McHugh TD, Chegou NN, et al. Tuberculosis: advances and challenges in development of new diagnostics and biomarkers. Lancet Infect Dis. 2018;18(7):e199-e210.
54. Colman RE, Mace A, Seifert M, Hetzel J, Mshaiel H, Suresh A, et al. Whole-genome and targeted sequencing of drug-resistant Mycobacterium tuberculosis on the iSeq100 and MiSeq: A performance, ease-of-use, and cost evaluation. PLoS Med. 2019;16(4):e1002794.

55. Hsiang E, Little KM, Haguma P, Hanrahan CF, Katamba A, Cattamanchi A, et al. Higher cost of implementing Xpert((R)) MTB/RIF in Ugandan peripheral settings: implications for cost-effectiveness. Int J Tuberc Lung Dis. 2016;20(9):1212-8.
56. Treatment Action Group. 2020 Pipeline Report: Tuberculosis Diagnostics. New York: Treatment Action Group; 2020 October 2020.
57. Detjen AK, DiNardo AR, Leyden J, Steingart KR, Menzies D, Schiller I, et al. Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis. Lancet Respir Med. 2015;3(6):451-61.
58. Central TB Division - Ministry of Health and Family Welfare; Government of India. Index-TB Guidelines: Guidelines on extra-pulmonary tuberculosis for India. New Delhi: All India Institute of Medical Sciences; 2016.
59. Denkinger CM, Schumacher SG, Gilpin C, Korobitsyn A, Wells WA, Pai M, et al. Guidance for the Evaluation of Tuberculosis Diagnostics That Meet the World Health Organization (WHO) Target Product Profiles: An Introduction to WHO Process and Study Design Principles. The Journal of infectious diseases. 2019;220(Supplement_3):S91-s8.
60. Regulation (EU) 2017/746 of the European Parliament and of the council of 5 April 2017 on in vitro diagnostic medical devices and repealing directive 98/79/EC and commission decision 2010/227/EU [Internet]. 2017. Available from: https://eur-lex.europa.eu/eli/reg/2017/746/oj.
61. World Health Organization. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update. Geneva: World Health Organization; 2019.
62. Ramsay A, Steingart KR, Cunningham J, Pai M. Translating tuberculosis research into global policies: the example of an international collaboration on diagnostics. Int J Tuberc Lung Dis. 2011;15(10):1283-93.
63. Vandenberg O, Martiny D, Rochas O, van Belkum A, Kozlakidis Z. Considerations for diagnostic COVID-19 tests. Nat Rev Microbiol. 2021;19(3):171-83.
64. Bossuyt PM. Testing COVID-19 tests faces methodological challenges. Journal of clinical epidemiology. 2020;126:172-6.
65. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. The Cochrane database of systematic reviews. 2021;3(3):Cd013705.
66. United States Food and Drug Administration. Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests - Guidance for Industry and FDA Staff. In: Services USDoHH, editor. Silver Spring, Maryland: Center for Devices and Radiological Health; 2007.
67. Karch A, Koch A, Zapf A, Zerr I, Karch A. Partial verification bias and incorporation bias affected accuracy estimates of diagnostic studies for biomarkers that were part of an existing composite gold standard. Journal of clinical epidemiology. 2016;78:73-82.
68. Alonzo TA, Pepe MS. Using a combination of reference tests to assess the accuracy of a new diagnostic test. Statistics in medicine. 1999;18(22):2987-3003.
69. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. Journal of clinical epidemiology. 2009;62(8):797-806.
70. Schiller I, van Smeden M, Hadgu A, Libman M, Reitsma JB, Dendukuri N. Bias due to composite reference standards in diagnostic accuracy studies. Statistics in medicine. 2016;35(9):1454-70.
71. Dendukuri N, Schiller I, de Groot J, Libman M, Moons K, Reitsma J, et al. Concerns about composite reference standards in diagnostic research. Bmj. 2018;360:j5779.
72. Walter SD, Macaskill P, Lord SJ, Irwig L. Effect of dependent errors in the assessment of diagnostic or screening test accuracy when the reference standard is imperfect. Statistics in medicine. 2012;31(11-12):1129-38.

73. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. Biometrics. 1985;41(4):959-68.
74. Hadgu A, Dendukuri N, Wang L. Evaluation of screening tests for detecting Chlamydia trachomatis: bias associated with the patient-infected-status algorithm. Epidemiology. 2012;23(1):72-82.
75. Walter SD, Irwig LM. Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. Journal of clinical epidemiology. 1988;41(9):923-37.
76. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard--a systematic review. Am J Epidemiol. 2014;179(4):423-31.
77. Hui SL, Zhou XH. Evaluation of diagnostic tests without gold standards. Statistical methods in medical research. 1998;7(4):354-70.
78. Collins J, Huynh M. Estimation of diagnostic test accuracy without full verification: a review of latent class methods. Statistics in medicine. 2014;33(24):4141-69.
79. Yang I, Becker MP. Latent variable modeling of diagnostic accuracy. Biometrics. 1997;53(3):948-58.
80. Dendukuri N, Hadgu A, Wang L. Modeling conditional dependence between diagnostic tests: a multiple latent variable model. Statistics in medicine. 2009;28(3):441-61.
81. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am J Epidemiol. 1995;141(3):263-72.
82. Wang Z, Dendukuri N, Joseph L. Understanding the effects of conditional dependence in research studies involving imperfect diagnostic tests. Statistics in medicine. 2017;36(3):466-80.
83. Qu Y, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. Biometrics. 1996;52(3):797-810.
84. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 3: Rapid diagnostics for tuberculosis detection. 2021 Update. Geneva: World Health Organization; 2021.
85. Treatment Action Group. 2021 Pipeline Report: Tuberculosis Diagnostics. New York: Treatment Action Group; 2021 November 2021.
86. World Health Organization. Target product profile for next-generation drug-susceptibility testing at peripheral centres. Geneva: World Health Organization; 2021.
87. World Health Organization. Catalogue of mutations in Mycobacterium tuberculosis complex and their association with drug resistance. In: World Health Organization, editor. Geneva: World Health Organization; 2021.
88. World Health Organization. Meeting report of the WHO expert consultation on the definition of extensively drug-resistant tuberculosis, 27-29 October 2020. Geneva: World Health Organization; 2021.
89. Penn-Nicholson A, Georghiou SB, Ciobanu N, Kazi M, Bhalla M, David A, et al. Detection of isoniazid, fluoroquinolone, ethionamide, amikacin, kanamycin, and capreomycin resistance by the Xpert MTB/XDR assay: a cross-sectional multicentre diagnostic accuracy study. Lancet Infect Dis. 2021.
90. Ruhwald M, editor The TB diagnostics pipeline during COVID-19. FIND and NDWG Annual Symposium; 2021 19 October 2021; Virtual: The Union.
91. SD Biosensor. STANDARD M10 [Webpage]. Suwon, Republic of Korea: SD Biosensor; 2021 [updated 2021.
92. Pai M, Schumacher SG, Abimbola S. Surrogate endpoints in global health research: still searching for killer apps and silver bullets? BMJ global health. 2018;3(2):e000755.
93. Schumacher SG, Sohn H, Qin ZZ, Gore G, Davis JL, Denkinger CM, et al. Impact of Molecular Diagnostics for Tuberculosis on Patient-Important Outcomes: A Systematic Review of Study Methodologies. PLoS One. 2016;11(3):e0151073.

94. Ochodo E, Kalema N, Schumacher S, Steingart K, Young T, Mallett S, et al. Variation in the observed effect of Xpert MTB/RIF testing for tuberculosis on mortality: A systematic review and analysis of trial design considerations [version 2; peer review: 3 approved]. Wellcome Open Research. 2020;4(173).
95. Ruhwald M, Carmona S, Pai M. Learning from COVID-19 to reimagine tuberculosis diagnosis. The Lancet Microbe. 2021.

8.2 Complete reference list

1. Regulation (EU) 2017/746 of the European Parliament and of the council of 5 April 2017 on in vitro diagnostic medical devices and repealing directive 98/79/EC and commission decision 2010/227/EU [Internet]. 2017. Available from: https://eur-lex.europa.eu/eli/reg/2017/746/oj.
2. Fiocruz establishes partnership with Chinese scientific institutions [press release]. São Paulo: Oswaldo Cruz Foundation, 26 February 2018 2018.
3. Albert H, Nathavitharana RR, Isaacs C, Pai M, Denkinger CM, Boehme CC. Development, roll-out and impact of Xpert MTB/RIF for tuberculosis: what lessons have we learnt and how can we do better? The European respiratory journal. 2016;48(2):516-25.
4. Albert PS, Dodd LE. A cautionary note on the robustness of latent class models for estimating diagnostic error without a gold standard. Biometrics. 2004;60(2):427-35.
5. Alonzo TA, Pepe MS. Using a combination of reference tests to assess the accuracy of a new diagnostic test. Statistics in medicine. 1999;18(22):2987-3003.
6. Amimo F, Lambert B, Magit A. What does the COVID-19 pandemic mean for HIV, tuberculosis, and malaria control? Tropical medicine and health. 2020;48:32.
7. Antel K, Oosthuizen J, Malherbe F, Louw VJ, Nicol MP, Maartens G, et al. Diagnostic accuracy of the Xpert MTB/Rif Ultra for tuberculosis adenitis. BMC infectious diseases. 2020;20(1):33-.
8. Armstrong GL, MacCannell DR, Taylor J, Carleton HA, Neuhaus EB, Bradbury RS, et al. Pathogen Genomics in Public Health. The New England journal of medicine. 2019;381(26):2569-80.
9. Auld AF, Fielding KL, Gupta-Wright A, Lawn SD. Xpert MTB/RIF - why the lack of morbidity and mortality impact in intervention trials? Transactions of the Royal Society of Tropical Medicine and Hygiene. 2016;110(8):432-44.
10. Bahr NC, Nuwagira E, Evans EE, Cresswell FV, Bystrom PV, Byamukama A, et al. Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: a prospective cohort study. The Lancet Infectious Diseases. 2018;18(1):68-75.
11. Banada PP, Naidoo U, Deshpande S, Karim F, Flynn JL, O'Malley M, et al. A novel sample processing method for rapid detection of tuberculosis in the stool of pediatric patients using the Xpert MTB/RIF assay. PLoS ONE. 2016;11 (3) (no pagination)(e0151980).
12. 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.
13. Bossuyt PM. Testing COVID-19 tests faces methodological challenges. Journal of clinical epidemiology. 2020;126:172-6.
14. Boyd R, Ford N, Padgen P, Cox H. Time to treatment for rifampicin-resistant tuberculosis: systematic review and meta-analysis. Int J Tuberc Lung Dis. 2017;21(11):1173-80.
15. Broder K, Babiker A, Myers C, White T, Jones H, Cardella J, et al. Test Agreement Between Roche Cobas 6800 and Cepheid GeneXpert Xpress SARS-CoV-2 Assays at High Cycle Threshold Ranges. J Clin Microbiol. 2020.
16. Butler-Laporte G, Lawandi A, Schiller I, Yao M, Dendukuri N, McDonald EG, et al. Comparison of Saliva and Nasopharyngeal Swab Nucleic Acid Amplification Testing for Detection of SARS-CoV-2: A Systematic Review and Meta-analysis. JAMA Internal Medicine. 2021;181(3):353-60.
17. Cabibbe AM, Battaglia S, Spitaleri A, Suresh A, Colman RE, Uplekar S, et al. A portable real-time solution for next generation sequencing-based diagnosis of drug-resistant tuberculosis direct from clinical samples. Int J Tuberc Lung Dis. 2019;23(19 Suppl 1):S268.

18. Campbell JR, Uppal A, Oxlade O, Fregonese F, Bastos ML, Lan Z, et al. Active testing of groups at increased risk of acquiring SARS-CoV-2 in Canada: costs and human resource needs. CMAJ : Canadian Medical Association journal = journal de l'Association medicale canadienne. 2020;192(40):E1146-E55.
19. Cazabon D, Pande T, Kik S, Van Gemert W, Sohn H, Denkinger C, et al. Market penetration of Xpert MTB/RIF in high tuberculosis burden countries: A trend analysis from 2014 - 2016. Gates Open Research. 2018;In press:1-8.
20. Central TB Division - Ministry of Health and Family Welfare; Government of India. Index-TB Guidelines: Guidelines on extra-pulmonary tuberculosis for India. New Delhi: All India Institute of Medical Sciences; 2016.
21. Cepheid Receives Emergency Use Authorization from FDA for Rapid SARS-CoV-2 Test [press release]. Sunnyvale, USA: Cepheid, 21 March 2020 2020.
22. Cepheid. Instructions for use: Xpert® Xpress SARS-CoV-2. In: Inc. C, editor. Sunnyvale, USA: Cepheid Inc.; 2020.
23. Chakravorty S, Simmons AM, Rowneki M, Parmar H, Cao Y, Ryan J, et al. The New Xpert MTB/RIF Ultra: Improving Detection of Mycobacterium tuberculosis and Resistance to Rifampin in an Assay Suitable for Point-of-Care Testing. mBio. 2017;8(4).
24. Chen H. VennDiagram: Generate High-Resolution Venn and Euler Plots. In: Boutros P, editor. 1.7.1 ed. Los Angeles: University of California Los Angeles; 2021. p. R package.
25. Chipinduro M, Mateveke K, Makamure B, Ferrand RA, Gomo E. Stool Xpert(R) MTB/RIF test for the diagnosis of childhood pulmonary tuberculosis at primary clinics in Zimbabwe. Int J Tuberc Lung Dis. 2017;21(2):161-6.
26. Christopher DJ, Schumacher SG, Michael JS, Luo R, Balamugesh T, Duraikannan P, et al. Performance of Xpert MTB/RIF on pleural tissue for the diagnosis of pleural tuberculosis. The European respiratory journal. 2013;42(5):1427-9.
27. Collins J, Huynh M. Estimation of diagnostic test accuracy without full verification: a review of latent class methods. Statistics in medicine. 2014;33(24):4141-69.
28. Colman RE, Mace A, Seifert M, Hetzel J, Mshaiel H, Suresh A, et al. Whole-genome and targeted sequencing of drug-resistant Mycobacterium tuberculosis on the iSeq100 and MiSeq: A performance, ease-of-use, and cost evaluation. PLoS Med. 2019;16(4):e1002794.
29. Comas I, Coscolla M, Luo T, Borrell S, Holt KE, Kato-Maeda M, et al. Out-of-Africa migration and Neolithic coexpansion of Mycobacterium tuberculosis with modern humans. Nature genetics. 2013;45(10):1176-82.
30. Conradie F, Diacon AH, Ngubane N, Howell P, Everitt D, Crook AM, et al. Treatment of Highly Drug-Resistant Pulmonary Tuberculosis. New England Journal of Medicine. 2020;382(10):893-902.
31. Cordova J, Shiloh R, Gilman RH, Sheen P, Martin L, Arenas F, et al. Evaluation of molecular tools for detection and drug susceptibility testing of Mycobacterium tuberculosis in stool specimens from patients with pulmonary tuberculosis. J Clin Microbiol. 2010;48(5):1820-6.
32. 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.
33. Cresswell FV, Tugume L, Bahr NC, Kwizera R, Bangdiwala AS, Musubire AK, et al. Xpert MTB/RIF Ultra for the diagnosis of HIV-associated tuberculous meningitis: a prospective validation study. Lancet Infect Dis. 2020.
34. Dabas H, Deo S, Sabharwal M, Pal A, Salim S, Nair L, et al. Initiative for Promoting Affordable and Quality Tuberculosis Tests (IPAQT): a market-shaping intervention in India. BMJ global health. 2019;4(6):e001539.
35. Das J, Kwan A, Daniels B, Satyanarayana S, Subbaraman R, Bergkvist S, et al. Use of standardised patients to assess quality of tuberculosis care: a pilot, cross-sectional study. Lancet Infect Dis. 2015;15(11):1305-13.

36. Dendukuri N, Hadgu A, Wang L. Modeling conditional dependence between diagnostic tests: a multiple latent variable model. Statistics in medicine. 2009;28(3):441-61.
37. Dendukuri N, Schiller I, de Groot J, Libman M, Moons K, Reitsma J, et al. Concerns about composite reference standards in diagnostic research. Bmj. 2018;360:j5779.
38. Deng S, Sun Y, Xia H, Liu Z, Gao L, Yang J, et al. Accuracy of Commercial Molecular Diagnostics for the Detection of Pulmonary Tuberculosis in China: A Systematic Review. Scientific reports. 2019;9(1):4553.
39. Denkinger CM, Dolinger D, Schito M, Wells W, Cobelens F, Pai M, et al. Target Product Profile of a Molecular Drug-Susceptibility Test for Use in Microscopy Centers. Journal of Infectious Diseases. 2015;211(suppl 2):S39-S49.
40. Denkinger CM, Kik SV, Cirillo DM, Casenghi M, Shinnick T, Weyer K, et al. Defining the Needs for Next Generation Assays for Tuberculosis. Journal of Infectious Diseases. 2015;211(suppl 2):S29-S38.
41. Denkinger CM, Schumacher SG, Gilpin C, Korobitsyn A, Wells WA, Pai M, et al. Guidance for the Evaluation of Tuberculosis Diagnostics That Meet the World Health Organization (WHO) Target Product Profiles: An Introduction to WHO Process and Study Design Principles. The Journal of infectious diseases. 2019;220(Supplement_3):S91-s8.
42. Detjen AK, DiNardo AR, Leyden J, Steingart KR, Menzies D, Schiller I, et al. Xpert MTB/RIF assay for the diagnosis of pulmonary tuberculosis in children: a systematic review and meta-analysis. Lancet Respir Med. 2015;3(6):451-61.
43. Dharmadhikari AS, Mphahlele M, Venter K, Stoltz A, Mathebula R, Masotla T, et al. Rapid impact of effective treatment on transmission of multidrug-resistant tuberculosis. Int J Tuberc Lung Dis. 2014;18(9):1019-25.
44. Di Tanna GL, Khaki AR, Theron G, McCarthy K, Cox H, Mupfumi L, et al. Effect of Xpert MTB/RIF on clinical outcomes in routine care settings: individual patient data meta-analysis. The Lancet Global health. 2019;7(2):e191-e9.
45. Dinnes J, Deeks JJ, Berhane S, Taylor M, Adriano A, Davenport C, et al. Rapid, point-of-care antigen and molecular-based tests for diagnosis of SARS-CoV-2 infection. The Cochrane database of systematic reviews. 2021;3(3):Cd013705.
46. Donald PR, Schaaf HS, Gie RP, Beyers N, Sirgel FA, Venter A. Stool microscopy and culture to assist the diagnosis of pulmonary tuberculosis in childhood. Journal of tropical pediatrics. 1996;42(5):311-2.
47. Donovan J, Thu DDA, Phu NH, Dung VTM, Quang TP, Nghia HDT, et al. Xpert MTB/RIF Ultra versus Xpert MTB/RIF for the diagnosis of tuberculous meningitis: a prospective, randomised, diagnostic accuracy study. The Lancet Infectious Diseases. 2020;20(3):299-307.
48. Drain PK, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, et al. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. Clin Microbiol Rev. 2018;31(4).
49. Dramé M, Tabue Teguo M, Proye E, Hequet F, Hentzien M, Kanagaratnam L, et al. Should RT-PCR be considered a gold standard in the diagnosis of COVID-19? Journal of Medical Virology. 2020;92(11):2312-3.
50. Dwamena B. MIDAS: Stata module for meta-analytical integration of diagnostic test accuracy studies. In: Economics BCDo, editor. Statistical Software Components2007.
51. El-Halfawy OM, Valvano MA. Antimicrobial Heteroresistance: an Emerging Field in Need of Clarity. Clinical Microbiology Reviews. 2015;28(1):191-207.
52. ElJilani MM, Abdusalam M, Abdalla A, Dalyoum T, Ahmad S, Alhudiri I, et al. Performance Evaluation of Seven Commercial Real-Time Polymerase Chain Reaction Assays for SARS-COV-2 Detection. Libyan J Med Sci. 2021;5(2):56-60.
53. FIND. Diagnostics Pipeline Tracker: Tuberculosis Geneva: FIND; 2020 [cited 2020 6 April, 2020]. Available from: https://www.finddx.org/dx-pipeline-status/.

54. New rapid molecular test for tuberculosis can simultaneously detect resistance to first- and second-line drugs [press release]. Geneva: FIND, 16 July 2020 2020.
55. Genoscreen. Deeplex(R) Myc-TB: From clinical samples to drug resistance profile. In: Genoscreen, editor. Genoscreen. Lille, France: Genoscreen; 2019.
56. Global Preparedness Monitoring Board. A World in Disorder: GPMB 2020 Annual Report. Geneva: World Health Organization; 2020.
57. Goodman LA. Exploratory latent structure analysis using both identifiable and unidentifiable models. Biometrika. 1974;61(2):215-31.
58. Graham SM, Cuevas LE, Jean-Philippe P, Browning R, Casenghi M, Detjen AK, et al. Clinical Case Definitions for Classification of Intrathoracic Tuberculosis in Children: An Update. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2015;61Suppl 3:S179-87.
59. Group TC-GS. Tuberculosis and COVID-19 co-infection: description of the global cohort. The European respiratory journal. 2021.
60. Guo Y, Zhou Y, Wang C, Zhu L, Wang S, Li Q, et al. Rapid, accurate determination of multidrug resistance in M. tuberculosis isolates and sputum using a biochip system. Int J Tuberc Lung Dis. 2009;13(7):914-20.
61. Hadgu A, Dendukuri N, Wang L. Evaluation of screening tests for detecting Chlamydia trachomatis: bias associated with the patient-infected-status algorithm. Epidemiology. 2012;23(1):72-82.
62. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. Journal of biomedical informatics. 2009;42(2):377-81.
63. Hasan Z, Shakoor S, Arif F, Mehnaz A, Akber A, Haider M, et al. Evaluation of Xpert MTB/RIF testing for rapid diagnosis of childhood pulmonary tuberculosis in children by Xpert MTB/RIF testing of stool samples in a low resource setting. BMC research notes. 2017;10(1):473.
64. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. Bmj. 2003;327(7414):557-60.
65. Hillemann D, Rüsch-Gerdes S, Boehme C, Richter E. Rapid molecular detection of extrapulmonary tuberculosis by the automated GeneXpert MTB/RIF system. Journal of clinical microbiology. 2011;49(4):1202-5.
66. Hofmann-Thiel S, Hoffman H. Evaluation of Fluorotype MTB for detection of Mycobacterium tuberculosis complex DNA in clinical specimens from a low-incidence country. BMC Infect Dis. 2014;14(59):1-5.
67. Horne DJ, Kohli M, Zifodya JS, Schiller I, Dendukuri N, Tollefson D, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. The Cochrane database of systematic reviews. 2019;6:Cd009593.
68. Hsiang E, Little KM, Haguma P, Hanrahan CF, Katamba A, Cattamanchi A, et al. Higher cost of implementing Xpert((R)) MTB/RIF in Ugandan peripheral settings: implications for cost-effectiveness. Int J Tuberc Lung Dis. 2016;20(9):1212-8.
69. Hui SL, Zhou XH. Evaluation of diagnostic tests without gold standards. Statistical methods in medical research. 1998;7(4):354-70.
70. India moving forward in research towards new tools for tuberculosis [press release]. New Delhi: Government of India, 2 November 2017 2017.
71. Jain D, Ghosh S, Teixeira L, Mukhopadhyay S. Pathology of pulmonary tuberculosis and non-tuberculous mycobacterial lung disease: Facts, misconceptions, and practical tips for pathologists. Seminars in diagnostic pathology. 2017;34(6):518-29.
72. Jekarl DW, Choi H, Kim JY, Lee S, Gweon TG, Lee HK, et al. Evaluating Diagnostic Tests for Helicobacter pylori Infection Without a Reference Standard: Use of Latent Class Analysis. Annals of laboratory medicine. 2020;40(1):68-71.

73. Jing H, Lu Z, Deng Y, Gao D, Li L, Graviss E, et al. Evaluation of Xpert MTB/RIF in detection of pulmonary and extrapulmonary tuberculosis cases in China. International Journal of Clinical and Experimental Pathology. 2017;10:4847-51.
74. Johns Hopkins University. COVID-19 dashboard by the center for systems science and engineering at Johns Hopkins University. In: University JH, editor. Coronavirus Research Centre. Baltimore, USA: Johns Hopkins University; 2021.
75. Joseph L, Gyorkos TW, Coupal L. Bayesian estimation of disease prevalence and the parameters of diagnostic tests in the absence of a gold standard. Am J Epidemiol. 1995;141(3):263-72.
76. Kantor HS, Poblete R, Pusateri SL. Nosocomial transmission of tuberculosis from unsuspected disease. The American journal of medicine. 1988;84(5):833-8.
77. Karch A, Koch A, Zapf A, Zerr I, Karch A. Partial verification bias and incorporation bias affected accuracy estimates of diagnostic studies for biomarkers that were part of an existing composite gold standard. Journal of clinical epidemiology. 2016;78:73-82.
78. Kendall EA, Shrestha S, Dowdy DW. The Epidemiological Importance of Subclinical Tuberculosis. A Critical Reappraisal. Am J Respir Crit Care Med. 2021;203(2):168-74.
79. Khobragade RN, Murthy N, Aloysius S, Surendran D, Rakesh PS, Balakrishnan S. Experience of integrated screening and testing for TB and COVID19 from Kerala, India. Public Health Pract (Oxf). 2021;2:100198-.
80. Kim YW, Kwak N, Seong MW, Kim EC, Yoo CG, Kim YW, et al. Accuracy of the Xpert® MTB/RIF assay for the diagnosis of extra-pulmonary tuberculosis in South Korea. Int J Tuberc Lung Dis. 2015;19(1):81-6.
81. Klinton JS, Heitkamp P, Rashid A, Faleye BO, Win Htat H, Hussain H, et al. One year of COVID-19 and its impact on private provider engagement for TB: A rapid assessment of intermediary NGOs in seven high TB burden countries. J Clin Tuberc Other Mycobact Dis. 2021;25:100277-.
82. Kohli M, Schiller I, Dendukuri N, Dheda K, Denkinger CM, Schumacher SG, et al. Xpert((R)) MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. The Cochrane database of systematic reviews. 2018;8:Cd012768.
83. Kumar MS, Surendran D, Manu MS, Rakesh PS, Balakrishnan S. Mortality due to TB-COVID-19 coinfection in India. Int J Tuberc Lung Dis. 2021;25(3):250-1.
84. Kwan A, Daniels B, Saria V, Satyanarayana S, Subbaraman R, McDowell A, et al. Variations in the quality of tuberculosis care in urban India: A cross-sectional, standardized patient study in two cities. PLoS medicine. 2018;15(9):e1002653-e.
85. LaCourse SM, Pavlinac PB, Cranmer LM, Njuguna IN, Mugo C, Gatimu J, et al. Stool Xpert MTB/RIF and urine lipoarabinomannan for the diagnosis of tuberculosis in hospitalized HIV-infected children. AIDS (London, England). 2018;32(1):69-78.
86. Lee J, Song JU. Diagnostic accuracy of the Cepheid Xpert Xpress and the Abbott ID NOW assay for rapid detection of SARS-CoV-2: A systematic review and meta-analysis. J Med Virol. 2021;93(7):4523-31.
87. Lee RA, Herigon JC, Benedetti A, Pollock NR, Denkinger CM. Performance of Saliva, Oropharyngeal Swabs, and Nasal Swabs for SARS-CoV-2 Molecular Detection: a Systematic Review and Meta-analysis. J Clin Microbiol. 2021;59(5).
88. Lee RS, Pai M. Real-Time Sequencing of Mycobacterium tuberculosis: Are We There Yet? J Clin Microbiol. 2017;55(5):1249-54.
89. Lee YC, Rogers JT, Rodriguez RM, Miller KD, Light RW. Adenosine deaminase levels in nontuberculous lymphocytic pleural effusions. Chest. 2001;120(2):356-61.
90. Lessells RJ, Cooke GS, McGrath N, Nicol MP, Newell ML, Godfrey-Faussett P. Impact of Point-of-Care Xpert MTB/RIF on Tuberculosis Treatment Initiation. A Cluster-randomized Trial. Am J Respir Crit Care Med. 2017;196(7):901-10.

91. Lewinsohn DM, Leonard MK, LoBue PA, Cohn DL, Daley CL, Desmond E, et al. Official American Thoracic Society/Infectious Diseases Society of America/Centers for Disease Control and Prevention Clinical Practice Guidelines: Diagnosis of Tuberculosis in Adults and Children. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2017;64(2):e1-e33.
92. Loveday M, Cox H, Evans D, Furin J, Ndjeka N, Osman M, et al. Opportunities from a new disease for an old threat: Extending COVID-19 efforts to address tuberculosis in South Africa. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde. 2020;110(12):1160-7.
93. Loyola S, Valle A, Montero S, Carrasco-Escobar G. Recomendaciones para describir de forma adecuada una curva epidemica de COVID-19. Rev Peru Med Exp Salud Publica. 2020;37(2):1-3.
94. Lusiba JK, Nakiyingi L, Kirenga BJ, Kiragga A, Lukande R, Nsereko M, et al. Evaluation of Cepheid's Xpert MTB/Rif test on pleural fluid in the diagnosis of pleural tuberculosis in a high prevalence HIV/TB setting. PLoS One. 2014;9(7):e102702.
95. Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Chapter 10: Analysing and Presenting Results. In: Collaboration TC, editor. Cochrane Handbok for Systematic Reviews of Diagnostic Test Accuracy. Version 1.0 ed: The Cochrane Collaboration; 2010.
96. MacLean E, Broger T, Yerlikaya S, Fernandez-Carballo BL, Pai M, Denkinger CM. A systematic review of biomarkers to detect active tuberculosis. Nat Microbiol. 2019;4(5):748-58.
97. Marais BJ, Gie RP, Hesseling AC, Schaaf HS, Lombard C, Enarson DA, et al. A refined symptom-based approach to diagnose pulmonary tuberculosis in children. Pediatrics. 2006;118(5):e1350-9.
98. Marcy O, Ung V, Goyet S, Borand L, Msellati P, Tejiokem M, et al. Performance of Xpert MTB/RIF and Alternative Specimen Collection Methods for the Diagnosis of Tuberculosis in HIV-Infected Children. Clinical Infectious Diseases. 2016;62(9):1161-8.
99. Mathur P, Sacks L, Auten G, Sall R, Levy C, Gordin F. Delayed diagnosis of pulmonary tuberculosis in city hospitals. Archives of internal medicine. 1994;154(3):306-10.
100. Maurer FP, Shubladze N, Kalmambetova G, Felker I, Kuchukhidze G, Drobniewski F, et al. Impact of the COVID-19 pandemic on tuberculosis national reference laboratory services in the WHO European Region, March to November 2020. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin. 2021;26(24):2100426.
101. McQuaid CF, Vassall A, Cohen T, Fiekert K, White RG. The impact of COVID-19 on TB: a review of the data. Int J Tuberc Lung Dis. 2021;25(6):436-46.
102. Meehan CJ, Goig GA, Kohl TA, Verboven L, Dippenaar A, Ezewudo M, et al. Whole genome sequencing of Mycobacterium tuberculosis: current standards and open issues. Nature reviews Microbiology. 2019;17(9):533-45.
103. Meldau R, Peter J, Theron G, Calligaro G, Allwood B, Symons G, et al. Comparison of same day diagnostic tools including Gene Xpert and unstimulated IFN-γ for the evaluation of pleural tuberculosis: a prospective cohort study. BMC pulmonary medicine. 2014;14:58.
104. Melsew YA, Doan TN, Gambhir M, Cheng AC, McBryde E, Trauer JM. Risk factors for infectiousness of patients with tuberculosis: a systematic review and meta-analysis. Epidemiol Infect. 2018;146(3):345-53.
105. Michot JM, Madec Y, Bulifon S, Thorette-Tcherniak C, Fortineau N, Noël N, et al. Adenosine deaminase is a useful biomarker to diagnose pleural tuberculosis in low to medium prevalence settings. Diagnostic microbiology and infectious disease. 2016;84(3):215-20.
106. Migliori GB, Nardell E, Yedilbayev A, D'Ambrosio L, Centis R, Tadolini M, et al. Reducing tuberculosis transmission: a consensus document from the World Health Organization Regional Office for Europe. The European respiratory journal. 2019;53(6).

107. Migliori GB, Thong PM, Akkerman O, Alffenaar JW, Álvarez-Navascués F, Assao-Neino MM, et al. Worldwide Effects of Coronavirus Disease Pandemic on Tuberculosis Services, January-April 2020. Emerging infectious diseases. 2020;26(11):2709-12.
108. Ministerio de Salud - el Perú. Avances en la Investigación y Diagnóstico de la Tuberculosis en el Perú. Lima: Ministerio de Salud - el Perú; 2018.
109. Norma téchnica de salud para la prevencion y control de la coninfeccion tuberculosis y virus de la immunodeficiencia humane en el Peru, Stat. 143-2018 (2 October 2018, 2018).
110. Mishra H, Reeve BWP, Palmer Z, Caldwell J, Dolby T, Naidoo CC, et al. Xpert MTB/RIF Ultra and Xpert MTB/RIF for diagnosis of tuberculosis in an HIV-endemic setting with a high burden of previous tuberculosis: a two-cohort diagnostic accuracy study. Lancet Respir Med. 2020.
111. 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, w64.
112. New Launch: Truenat(TM) SARS CoV-2 Chip-based Real Time PCR Test for COVID-19 [press release]. Goa: Molbio Diagnostics Private Limited,2020.
113. Moore NM, Li H, Schejbal D, Lindsley J, Hayden MK. Comparison of two commercial molecular tests and a laboratory-developed modification of the CDC 2019-nCoV RT-PCR assay for the detection of SARS-CoV-2. J Clin Microbiol. 2020.
114. Moussa H, Bayoumi FS, Mohamed AM. Gene Xpert for Direct Detection of Mycobacterium Tuberculosis in Stool Specimens from Children with Presumptive Pulmonary Tuberculosis. Annals of clinical and laboratory science. 2016;46(2):198-203.
115. MSF. Step up for TB 2020: Tuberculosis Policies in 37 countries. Geneva: Medicins Sans Frontieres; 2020 November 2020.
116. Nair D, Navneethapandian PD, Tripathy JP, Harries AD, Klinton JS, Watson B, et al. Impact of rapid molecular diagnostic tests on time to treatment initiation and outcomes in patients with multidrug-resistant tuberculosis, Tamil Nadu, India. Transactions of the Royal Society of Tropical Medicine and Hygiene. 2016;110(9):534-41.
117. National Institute for Communicable Diseases. National Institute for Communicable Diseases Annual Overview 2018/2019. Capetown: National Institute for Communicable Diseases; 2018 6 April, 2020.
118. Nicol MP, Spiers K, Workman L, Isaacs W, Munro J, Black F, et al. Xpert MTB/RIF testing of stool samples for the diagnosis of pulmonary tuberculosis in children. Clinical Infectious Diseases. 2013;57(3):e18-e21.
119. Nikam C, Jagannath M, Narayanan MM, Ramanabhiraman V, Kazi M, Shetty A, et al. Rapid diagnosis of Mycobacterium tuberculosis with Truenat MTB: a near-care approach. PLoS One. 2013;8(1):e51121.
120. Ochodo E, Kalema N, Schumacher S, Steingart K, Young T, Mallett S, et al. Variation in the observed effect of Xpert MTB/RIF testing for tuberculosis on mortality: A systematic review and analysis of trial design considerations [version 2; peer review: 3 approved]. Wellcome Open Research. 2020;4(173).
121. Orikiriza P, Nansumba M, Nyehangane D, Bastard M, Mugisha IT, Nansera D, et al. Xpert MTB/RIF diagnosis of childhood tuberculosis from sputum and stool samples in a high TB-HIV-prevalent setting. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology. 2018.
122. Ortega C, Wood R, Murton H, Andama A, Cattamanchi A, Dixon R, et al. Diagnosis of pulmonary tuberculosis by oral swab analysis (OSA): optimisation and development of non-sputum, point-of-care methods. Int J Tuberc Lung Dis. 2019;23(10 Suppl 1):S211.
123. Otero L, Shah L, Verdonck K, Battaglioli T, Brewer T, Gotuzzo E, et al. A prospective longitudinal study of tuberculosis among household contacts of smear-positive tuberculosis cases in Lima, Peru. BMC infectious diseases. 2016;16:259-.

124. Oxlade O, Sugarman J, Alvarez GG, Pai M, Schwartzman K. Xpert(R)MTB/RIF for the Diagnosis of Tuberculosis in a Remote Arctic Setting: Impact on Cost and Time to Treatment Initiation. PLoS One. 2016;11(3):e0150119.
125. Pai M. Global health technologies: time to re-think the 'trickle down' model. Forbes [Internet]. 2020 18 February 2020. Available from: https://www.forbes.com/sites/madhukarpai/2020/02/17/global-health-technologies-time-to-re-think-the-trickle-down-model/#6eed3b6344d9.
126. Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. Nature reviews Disease primers. 2016;2:16076.
127. Pai M, Dendukuri N, Wang L, Joshi R, Kalantri S, Rieder HL. Improving the estimation of tuberculosis infection prevalence using T-cell-based assay and mixture models. Int J Tuberc Lung Dis. 2008;12(8):895-902.
128. Pai M, Denkinger CM, Kik SV, Rangaka MX, Zwerling A, Oxlade O, et al. Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. Clin Microbiol Rev. 2014;27(1):3-20.
129. Pai M, Furin J. Tuberculosis innovations mean little if they cannot save lives. eLife. 2017;6.
130. Pai M, Schumacher SG, Abimbola S. Surrogate endpoints in global health research: still searching for killer apps and silver bullets? BMJ global health. 2018;3(2):e000755.
131. Pang Y, Dong H, Tan Y, Deng Y, Cai X, Jing H, et al. Rapid diagnosis of MDR and XDR tuberculosis with the MeltPro TB assay in China. Scientific reports. 2016;6:25330-.
132. Pašović M, Leach-Kemon K, Troeger C, Vos T, Lozano R. Countries Hit Hardest by COVID-19 Washinton DC: Think Global Health; 2021 [cited 2021 23 November]. Available from: https://www.thinkglobalhealth.org/article/countries-hit-hardest-covid-19.
133. Pearson JD, Trcka D, Lu S, Hyduk SJ, Jen M, Aynaud MM, et al. Comparison of SARS-CoV-2 indirect and direct RT-qPCR detection methods. Virology journal. 2021;18(1):99.
134. Penn-Nicholson A, Georghiou SB, Ciobanu N, Kazi M, Bhalla M, David A, et al. Detection of isoniazid, fluoroquinolone, ethionamide, amikacin, kanamycin, and capreomycin resistance by the Xpert MTB/XDR assay: a cross-sectional multicentre diagnostic accuracy study. Lancet Infect Dis. 2021.
135. Pepe MS. The Statistical Evaluation of Medical Tests for Classification and Prediction. 1 ed. Oxford: Oxford University Press; 2003.
136. Pereira GR, Barbosa MS, Dias NJD, Almeida CPB, Silva DR. Impact of introduction of Xpert MTB/RIF test on tuberculosis (TB) diagnosis in a city with high TB incidence in Brazil. PLoS One. 2018;13(3):e0193988.
137. Qu Y, Tan M, Kutner MH. Random effects models in latent class analysis for evaluating accuracy of diagnostic tests. Biometrics. 1996;52(3):797-810.
138. Ramsay A, Steingart KR, Cunningham J, Pai M. Translating tuberculosis research into global policies: the example of an international collaboration on diagnostics. Int J Tuberc Lung Dis. 2011;15(10):1283-93.
139. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. Journal of clinical epidemiology. 2005;58(10):982-90.
140. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. Journal of clinical epidemiology. 2009;62(8):797-806.
141. Revised National Tuberculosis Control Programme. National Strategic Plan for Tuberculosis Elimination 2017–2025. New Delhi: Directorate General of Health Services; 2017.
142. Riley RL, Mills CC, O'Grady F, Sultan LU, Wittstadt F, Shivpuri DN. Infectiousness of air from a tuberculosis ward. Ultraviolet irradiation of infected air: comparative infectiousness of different patients. The American review of respiratory disease. 1962;85:511-25.

143. Rocchetti TT, Silbert S, Gostnell A, Kubasek C, Widen R. Validation of a Multiplex Real-Time PCR Assay for Detection of Mycobacterium spp., Mycobacterium tuberculosis Complex, and Mycobacterium avium Complex Directly from Clinical Samples by Use of the BD Max Open System. J Clin Microbiol. 2016;54(6):1644-7.
144. Roche Molecular Systems I. Cobas(R) MTB: Nucleic acid test for use on the Cobas(R) 6800/8800 systems. Basel: Roche Molecular Systems; 2018.
145. Rouillon A, Perdrizet S, Parrot R. Transmission of tubercle bacilli: The effects of chemotherapy. Tubercle. 1976;57(4):275-99.
146. RStudio Team. RStudio: Integrated Development for R. In: Rstudio I, editor. Boston: RStudio, Inc.; 2016.
147. Ruhwald M, editor The TB diagnostics pipeline during COVID-19. FIND and NDWG Annual Symposium; 2021 19 October 2021; Virtual: The Union.
148. Ruhwald M, Carmona S, Pai M. Learning from COVID-19 to reimagine tuberculosis diagnosis. The Lancet Microbe. 2021.
149. Schiller I, van Smeden M, Hadgu A, Libman M, Reitsma JB, Dendukuri N. Bias due to composite reference standards in diagnostic accuracy studies. Statistics in medicine. 2016;35(9):1454-70.
150. Schofield MR, Maze MJ, Crump JA, Rubach MP, Galloway R, Sharples KJ. On the robustness of latent class models for diagnostic testing with no gold standard. Statistics in medicine. 2021;40(22):4751-63.
151. Schumacher SG, Denkinger CM. The impact of Xpert MTB/RIF-do we have a final answer? The Lancet Global health. 2019;7(2):e161-e2.
152. Schumacher SG, Sohn H, Qin ZZ, Gore G, Davis JL, Denkinger CM, et al. Impact of Molecular Diagnostics for Tuberculosis on Patient-Important Outcomes: A Systematic Review of Study Methodologies. PLoS One. 2016;11(3):e0151073.
153. Schumacher SG, van Smeden M, Dendukuri N, Joseph L, Nicol MP, Pai M, et al. Diagnostic Test Accuracy in Childhood Pulmonary Tuberculosis: A Bayesian Latent Class Analysis. Am J Epidemiol. 2016;184(9):690-700.
154. SD Biosensor. STANDARD M10 [Webpage]. Suwon, Republic of Korea: SD Biosensor; 2021 [updated 2021.
155. Seddon JA, Tugume L, Solomons R, Prasad K, Bahr NC. The current global situation for tuberculous meningitis: epidemiology, diagnostics, treatment and outcomes. Wellcome Open Res. 2019;4:167.
156. Shah M, Paradis S, Betz J, Beylis N, Bharadwaj R, Caceres T, et al. Multicenter Study of the Accuracy of the BD MAX MDR-TB Assay for Detection of Mycobacterium tuberculosis Complex and Mutations Associated with Resistance to Rifampin and Isoniazid. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2019.
157. Sharma SK, Kohli M, Chaubey J, Yadav RN, Sharma R, Singh BK, et al. Evaluation of Xpert MTB/RIF assay performance in diagnosing extrapulmonary tuberculosis among adults in a tertiary care centre in India. The European respiratory journal. 2014;44(4):1090-3.
158. Shea J, Halse TA, Lapierre P, Shudt M, Kohlerschmidt D, Van Roey P, et al. Comprehensive Whole-Genome Sequencing and Reporting of Drug Resistance Profiles on Clinical Cases of Mycobacterium tuberculosis in New York State. J Clin Microbiol. 2017;55(6):1871-82.
159. Shete PB, Farr K, Strnad L, Gray CM, Cattamanchi A. Diagnostic accuracy of TB-LAMP for pulmonary tuberculosis: a systematic review and meta-analysis. BMC infectious diseases. 2019;19(1):268-.
160. Singhroy DN, MacLean E, Kohli M, Lessem E, Branigan D, England K, et al. Adoption and uptake of the lateral flow urine LAM test in countries with high tuberculosis and HIV/AIDS burden: current landscape and barriers [version 1; peer review: 2 approved]. Gates Open Res. 2020;4(24).

161. Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, Cunningham J, et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. Lancet Infect Dis. 2006;6(9):570-81.
162. Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert(R) MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. The Cochrane database of systematic reviews. 2014(1):Cd009593.
163. Stevenson M. Tools for the Analysis of Epidemiological Data. In: Stevenson M, editor. 2.0.33 ed. Melbourne: The University of Melbourne; 2017.
164. Stop TB Partnership. The potential impact of the COVID-19 response on tuberculosis in high-burden countries: a modelling analysis. Geneva: Stop TB Partnership; 2020.
165. Stop TB Partnership. The Impact of COVID-19 on the TB Epidemic: A Community Perspective. Geneva: Stop TB Partnership; 2020.
166. Stop TB Partnership. 12 Months of COVID-19 Eliminated 12 Years of Progress in the Global Fight Against Tuberculosis. Geneva: Stop TB Partnership; 2021 18 March 2021. Contract No.: 2.
167. Stop TB Partnership. Global Drug Facility: Diagnostics Catalogue. Geneva: Stop TB Partnership; 2021.
168. Stop TB Partnership, McGill International TB Centre, Arcady Group. Re-Imagining TB Care Geneva: Stop TB Partnership; 2020 [
169. Storla DG, Yimer S, Bjune GA. A systematic review of delay in the diagnosis and treatment of tuberculosis. BMC Public Health. 2008;8:15.
170. Subbaraman R, Nathavitharana RR, Mayer KH, Satyanarayana S, Chadha VK, Arinaminpathy N, et al. Constructing care cascades for active tuberculosis: A strategy for program monitoring and identifying gaps in quality of care. PLoS Med. 2019;16(2):e1002754.
171. Subbaraman R, Nathavitharana RR, Satyanarayana S, Pai M, Thomas BE, Chadha VK, et al. The Tuberculosis Cascade of Care in India's Public Sector: A Systematic Review and Meta-analysis. PLoS Med. 2016;13(10):e1002149.
172. Sun Y, Gao L, Xia H, Yang Z, Deng S, Yang J, et al. Accuracy of molecular diagnostic tests for drug-resistant tuberculosis detection in China: a systematic review. Int J Tuberc Lung Dis. 2019;23(8):931-42.
173. Tang N, Frank A, Pahalawatta V, Lampinen J, Coblenz-Korte A, Dunn C, et al. Analytical and clinical performance of Abbott RealTime MTB, an assay for detection of Mycobacterium tuberculosis in pulmonary specimens. Tuberculosis (Edinburgh, Scotland). 2015;95:613-9.
174. The Global Fund. Briefing Note: Testing for both Tuberculosis and SARS-CoV-2. Geneva: The Global Fund; 2021 13 October 2021.
175. The TB/COVID-19 Global Study Group. Tuberculosis and COVID-19 co-infection: description of the global cohort. The European respiratory journal. 2021;in press.
176. 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. Lancet. 2014;383(9915):424-35.
177. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJ. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. PLoS One. 2011;6(4):e17601.
178. Togun T, Kampmann B, Stoker NG, Lipman M. Anticipating the impact of the COVID-19 pandemic on TB patients and TB control programmes. Annals of clinical microbiology and antimicrobials. 2020;19(1):21.
179. Treatment Action Group. 2019 Pipeline Report: Tuberculosis Diagnostics. New York: Treatment Action Group; 2019.
180. Treatment Action Group. 2020 Pipeline Report: Tuberculosis Diagnostics. New York: Treatment Action Group; 2020 October 2020.

181. Treatment Action Group. 2021 Pipeline Report: Tuberculosis Diagnostics. New York: Treatment Action Group; 2021 November 2021.
182. Ugarte-Gil C. TB Case Detection Gap 2020-2019 in Peru. In: MacLean E, editor. Figure ed. Lima: Ugarte-Gil, C.; 2021. p. 1.
183. Ugarte-Gil C, Curisinche M, Figueroa CJ, Gotuzzo E, Rios R. COVID-19 among tuberculosis patients in Peru: An operational report from the national registry. 52nd Union World Conference on Lung Health; 20 October 2021; Virtual: The Union for Tuberculosis and Lung Health; 2021.
184. Unitaid. UNITAID end-of-project evaluation: TB GeneXpert – Scaling up access to contemporary diagnostics for TB. Geneva: Unitaid; 2017.
185. United States Food and Drug Administration. Statistical Guidance on Reporting Results from Studies Evaluating Diagnostic Tests - Guidance for Industry and FDA Staff. In: Services USDoHH, editor. Silver Spring, Maryland: Center for Devices and Radiological Health; 2007.
186. Coronavirus (COVID-19) Update: FDA Issues first Emergency Use Authorization for Point of Care Diagnostic [press release]. Atlanta: United States Food and Drug Administration, 21 March, 2020. 2020.
187. Urbanczik R, Rieder HL. Scaling up tuberculosis culture services: a precautionary note. Int J Tuberc Lung Dis. 2009;13(7):799-800.
188. USAID, Stop TB Partnership. Simultaneous, integrated diagnostic testing approach to detect COVID-19 and TB in high TB burden countries. Geneva: Stop TB Partnership; 2021.
189. Vacek PM. The effect of conditional dependence on the evaluation of diagnostic tests. Biometrics. 1985;41(4):959-68.
190. van Smeden M, Naaktgeboren CA, Reitsma JB, Moons KG, de Groot JA. Latent class models in diagnostic studies when there is no reference standard--a systematic review. Am J Epidemiol. 2014;179(4):423-31.
191. Vandenberg O, Martiny D, Rochas O, van Belkum A, Kozlakidis Z. Considerations for diagnostic COVID-19 tests. Nat Rev Microbiol. 2021;19(3):171-83.
192. Venkatesan P. COVID-19 diagnostics—not at the expense of other diseases. The Lancet Microbe. 2020;1(2):e64.
193. Walker TM, Kohl TA, Omar SV, Hedge J, Del Ojo Elias C, Bradley P, et al. Whole-genome sequencing for prediction of Mycobacterium tuberculosis drug susceptibility and resistance: a retrospective cohort study. The Lancet Infectious Diseases. 2015;15(10):1193-202.
194. Walter SD, Irwig LM. Estimation of test error rates, disease prevalence and relative risk from misclassified data: a review. Journal of clinical epidemiology. 1988;41(9):923-37.
195. Walter SD, Macaskill P, Lord SJ, Irwig L. Effect of dependent errors in the assessment of diagnostic or screening test accuracy when the reference standard is imperfect. Statistics in medicine. 2012;31(11-12):1129-38.
196. Walters E, Scott L, Nabeta P, Demers AM, Reubenson G, Bosch C, et al. Molecular detection of Mycobacterium tuberculosis from stool in young children using a novel centrifugation-free processing method. J Clin Microbiol. 2018.
197. Walters E, Van Der Zalm MM, Palmer M, Bosch C, Demers AM, Draper H, et al. Xpert MTB/RIF on Stool Is Useful for the Rapid Diagnosis of Tuberculosis in Young Children with Severe Pulmonary Disease. Pediatric Infectious Disease Journal. 2017;36(9):837-43.
198. Walzl G, McNerney R, du Plessis N, Bates M, McHugh TD, Chegou NN, et al. Tuberculosis: advances and challenges in development of new diagnostics and biomarkers. Lancet Infect Dis. 2018;18(7):e199-e210.
199. Wang G, Wang S, Yang X, Sun Q, Jiang G, Huang M, et al. Accuracy of Xpert MTB/RIF Ultra for the Diagnosis of Pleural TB in a Multicenter Cohort Study. Chest. 2020;157(2):268-75.
200. Wang Z, Dendukuri N, Joseph L. Understanding the effects of conditional dependence in research studies involving imperfect diagnostic tests. Statistics in medicine. 2017;36(3):466-80.

201. Wen H, Li P, Ma H, Lv G. Diagnostic accuracy of Xpert MTB/RIF assay for musculoskeletal tuberculosis: a meta-analysis. Infection and drug resistance. 2017;10:299-305.
202. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Annals of internal medicine. 2011;155(8):529-36.
203. Wolters F, van de Bovenkamp J, van den Bosch B, van den Brink S, Broeders M, Chung NH, et al. Multi-center evaluation of cepheid xpert® xpress SARS-CoV-2 point-of-care test during the SARS-CoV-2 pandemic. Journal of clinical virology : the official publication of the Pan American Society for Clinical Virology. 2020;128:104426-.
204. World Health Organization. Molecular Line Probe Assays for Rapid Screening of Patients at Risk of Multi-Drug Resistant Tuberculosis: Policy Statement. Geneva: World Health Organization; 2008.
205. WHO endorses new rapid tuberculosis test [press release]. London | Geneva: World Health Organization, 8 December 2010 2010.
206. World Health Organization. Policy Update. Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF assay for the diagnosis of pulmonary and extrapulmonary TB in adults and children. Geneva, Switzerland: World Health Organization; 2013.
207. World Health Organization. WHO End TB Strategy. Geneva: World Health Organization; 2015.
208. World Health Organization. Policy framework: Implementing tuberculosis diagnostics. Geneva: World Health Organization; 2015.
209. World Health Organization. The use of loop-mediated isothermal amplification (TB-LAMP) for the diagnosis of pulmonary tuberculosis: Policy guidance. Geneva: World Health Organisation; 2016.
210. World Health Organization. The use of the Xpert MTB/RIF assay for the diagnosis of TB. Meeting Report. Geneva, Switzerland: World Health Organization; 2016.
211. World Health Organization. The use of molecular line probe assays for the detection of resistance to second-line anti-tuberculosis drugs. Geneva: World Health Organization; 2016.
212. Next-generation Xpert® MTB/RIF Ultra assay recommended by WHO [press release]. Geneva: World Health Organization2017.
213. World Health Organization. Global Tuberculosis Report 2018. Geneva, Switzerland: World Health Organization; 2018.
214. World Health Organization. Technical guide on next-generation sequencing technologies for the detection of mutations associated with drug resistance in Mycobacterium tuberculosis complex. Geneva: World Health Organization; 2018.
215. World Health Organization. Global tuberculosis report 2019. Geneva: World Health Organization; 2019.
216. World Health Organization. Lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis of active tuberculosis in people living with HIV. Policy update. Geneva: World Health Organization; 2019.
217. World Health Organization. Relational sequencing TB data platform [website]. Geneva: World Health Organization; 2019.
218. World Health Organization. WHO consolidated guidelines on drug-resistant tuberculosis treatment. Geneva: World Health Organization; 2019.
219. World Health Organization. Evaluation of centralized assays for TB detection and detection of resistance to rifampicin and isoniazid: WHO Technical Expert Consultation Report. Geneva: World Health Organization; 2019. Contract No.: WHO-CDS-TB-2019.14.
220. World Health Organization. Molecular assays intended as initial tests for the diagnosis of pulmonary and extrapulmonary TB and rifampicin resistance in adults and children: rapid communication. Policy update. Geneva: World Health Organization; 2020.

221. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 3: Rapid diagnostics for tuberculosis detection. Geneva: World Health Organization; 2020.
222. World Health Organization. Global Tuberculosis Report 2020. Geneva: World Health Organization; 2020.
223. World Health Organization. WHO consolidated guidelines on tuberculosis: Module 4: Treatment: drug-resistant tuberculosis treatment. Geneva: World Health Organization; 2020.
224. World Health Organization. WHO global lists of high burden countries for tuberculosis (TB), TB/HIV and multidrug/rifampicin-resistant TB (MDR/RR-TB), 2021–2025. Geneva: World Health Organization; 2021.
225. World Health Organization. Target product profile for next-generation drug-susceptibility testing at peripheral centres. Geneva: World Health Organization; 2021.
226. World Health Organization. Global Tuberculosis Report 2021. Geneva: World Health Organization; 2021.
227. World Health Organization. Meeting report of the WHO expert consultation on the definition of extensively drug-resistant tuberculosis, 27-29 October 2020. Geneva: World Health Organization; 2021.
228. Tuberculosis profile: India [Internet]. World Health Organization. 2021 [cited 28 November 2021]. Available from: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&entity_type=%22country%22&lan=%22EN%22&iso2=%22IN%22.
229. Tuberculosis profile: Peru [Internet]. World Health Organization. 2021 [cited 28 November 2021]. Available from: https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&entity_type=%22country%22&lan=%22EN%22&iso2=%22PE%22.
230. World Health Organization. WHO consolidated guidelines on tuberculosis. Module 3: Rapid diagnostics for tuberculosis detection. 2021 Update. Geneva: World Health Organization; 2021.
231. Worldometers.info. Worldometer: Peru dashboard Dover, Delware, United States2021 [updated 6 December 2021; cited 2021 6 December]. Dashboard]. Available from: https://www.worldometers.info/coronavirus/country/peru/.
232. Xie YL, Chakravorty S, Armstrong DT, Hall SL, Via LE, Song T, et al. Evaluation of a Rapid Molecular Drug-Susceptibility Test for Tuberculosis. The New England journal of medicine. 2017;377(11):1043-54.
233. Yan L, Tang S, Yang Y, Shi X, Ge Y, Sun W, et al. A Large Cohort Study on the Clinical Value of Simultaneous Amplification and Testing for the Diagnosis of Pulmonary Tuberculosis. Medicine. 2016;95(4):e2597.
234. Yang I, Becker MP. Latent variable modeling of diagnostic accuracy. Biometrics. 1997;53(3):948-58.
235. Yates TA, Khan PY, Knight GM, Taylor JG, McHugh TD, Lipman M, et al. The transmission of Mycobacterium tuberculosis in high burden settings. Lancet Infect Dis. 2016;16(2):227-38.
236. TGen and ABL plan global rollout of advanced test for TB, one of the world's most deadly pathogens [press release]. Flaggstaff, Arizona and Luxembourg City, Luxembourg: TGen, 31 July 2019 2019.
237. Zhen W, Smith E, Manji R, Schron D, Berry GJ. Clinical Evaluation of Three Sample-to-Answer Platforms for Detection of SARS-CoV-2. J Clin Microbiol. 2020;58(8).
238. Zignol M, Cabibbe AM, Dean AS, Glaziou P, Alikhanova N, Ama C, et al. Genetic sequencing for surveillance of drug resistance in tuberculosis in highly endemic countries: a multi-country population-based surveillance study. Lancet Infect Dis. 2018;18(6):675-83.

8.3 Doctoral training publication list

1. Huddart S*, **MacLean E***, Pai M. Location, location, location: tuberculosis services in highest burden countries. Lancet Glob Health. 2016 Dec;4(12):e907-e908.

*co-first authors

2. **MacLean E**, Huddart S, Pai M. Molecular diagnosis of tuberculosis: we need solutions that span the healthcare value chain. Expert Rev Mol Diagn. 2017 Jan;17(1):5-7.

3. **MacLean E**, Broger T. A 10-Gene Signature for the Diagnosis and Treatment Monitoring of Active Tuberculosis Using a Molecular Interaction Network Approach. EBioMed. 2017 Feb;16:22-23.

4. Yerlikaya S, Broger T, **MacLean E**, Pai M, Denkinger CM. A tuberculosis biomarker database: The key to novel TB diagnostics, Int J Inf Dis. 2017 Mar;56:253-257.

5. **MacLean E**, Pai M. Urine lipoarabinomannan for tuberculosis diagnosis: evolution and prospects. J Clin Chem. 2018 Aug;64(8):1133-1135.

6. Togun Y, **MacLean E**, Kampmann B, Pai M. Biomarkers for diagnosis of childhood tuberculosis: a systematic review. PLoS One. 2018 Sep 13;13(9):e0204029.

7. Soucy JR, Kutcher SA, **MacLean E**, Sewitch MJ. You only die twice. Am J Public Health. 2019 Jan;109(1):e9.

8. **MacLean E**, Saravu K, Pai M. Diagnosing active tuberculosis in people living with HIV: an on-going challenge. Curr Opin HIV/AIDS. 2019 Jan;14(1):46-54.

9. **MacLean E***, Broger T*, Yerlikaya S, Fernandez-Carballo BL, Pai M, Denkinger CM. A systematic review of biomarkers to detect active tuberculosis. Nature Microbiol. 2019 May;4(5):748-758.

10. **MacLean E**, Sulis G, Denkinger CM, Johnston JC, Pai M, Ahmad Khan F. Diagnostic accuracy of stool Xpert MTB/RIF for the detection of pulmonary tuberculosis in children: a systematic review and meta-analysis. J Clin Microbiol. 2019 May 24;57(6):e02057-18.

11. Singhroy DN, **MacLean E**, Kohli M, Lessem E, Branigan D, England K, Suleiman K, Drain PK, Ruhwald M, Schumacher S, Denkinger CM, Waning B, van Gemert W, Pai M. Adoption and uptake of the lateral flow urine LAM test in countries with high tuberculosis and HIV/AIDS burden: current landscape and barriers. Gates Open Res. 2020 Apr 7;4:24.

12. Bastos, ML, Tavaziva, G, Kunal Abidi S, Campbell JR, Haraoui L-P, Johnston JC, Lan Z, Law S, **MacLean E**, Trajman A, Menzies D, Benedetti A, Ahmad Khan F. Diagnostic accuracy of serological tests for COVID-19: a systematic review and meta-analysis. BMJ. 2020 Jul 1;370:m2516.

13. **MacLean E**, Kohli M, Weber SF, Suresh A, Schumacher SG, Denkinger CM, Pai M. Advances in molecular diagnosis of tuberculosis. J Clin Microbiol. 2020 Sep 22;58(10):e01582-19.

14. **MacLean E***, Bigio J*, Singh U*, Klinton JS, Pai M. Global tuberculosis awards must do better with equity, diversity, and inclusion. Lancet. 2021 Jan 16;397(10270):192-3.

15. Kohli M, **MacLean E**, Pai M, Denkinger CM, Schumacher SG. Diagnostic accuracy of centralized assays for TB detection and detection of resistance to rifampicin and isoniazid: A systematic review and meta-analysis. Eur Resp J. 2021 Feb 4;57(2):2000747.

16. Fernandez-Carballo BL, Escadafal C, **MacLean E**, Kapasi AJ, Dittrich S. Distinguishing bacterial versus non-bacterial causes of febrile illness – a systematic review of host biomarkers. J Infect. 2021 Apr;82(4):1-10.

17. Bigio J, Kohli M, Klinton JS, **MacLean E**, Gore G, Small PM, Ruhwald M, Weber SF, Jha S, Pai M. Diagnostic accuracy of point-of-care ultrasound for pulmonary tuberculosis: a systematic review. PLoS One. 2021 May 7;16(5):e0251236.

18. **MacLean E**, Nathavitharana RR. Progress towards developing sensitive non-sputum based TB diagnostic tests: the promise of urine cell-free DNA. J Clin Microbiol. 2021 Jul 19; 59(8):e0070621.

19. Heirwegh E, **MacLean E**, He L, Kamhawi S, Sagan SM, Olivier M. Sandfly Fever Sicilian Virus-Leishmania major co-infection modulates innate inflammatory response favoring myeloid cell infections and skin hyperinflammation. PLoS Negl Trop Dis. 2021 Jul 26;15(7):e0009638.

20. **MacLean E**, Dendukuri N. Latent class analysis and the need for clear reporting of methodology. Clin Inf Dis. 2021 Oct 5;73(7):e2285-e2286.

21. Zimmer AJ, Lainati F, Aguilera Vasquez N, Chedid C, McGrath S, Benedetti A, **MacLean E**, Ruhwald M, Denkinger CM, Kohli M. Biomarkers for active pulmonary tuberculosis treatment response: a systematic review. J Clin Microbiol. 2022 Feb 16:60(2);e0185921.

23. **MacLean EL**, Villa-Castillo L, Ruhwald M, Ugarte-Gil C, Pai M. Commentary: Integrated testing for TB and COVID-19. Med. 2022 Mar 11;3(3):162-6.

24. Bigio J, **MacLean E**, Aguilera Vasquez N, Kohli M, Gore G, Pai M, Adam P. Most common reasons for primary care visits in low- and middle-income countries: a systematic review. PLOS Glob Public Health. 2022. Accepted.

24. **MacLean E**, Kohli M, Köppel L, Schiller I, Sharma SK, Pai M, Denkinger CM, Dendukuri, N. Bayesian latent class analysis versus composite reference standards for extra-pulmonary tuberculosis diagnostic tests. *Diagnos and Prognos Res*. 2022. Accepted.

25. Anand A*, Bigio J*, **MacLean E***, Underwood T, Pai N, Carmona S, Schumacher SG, Toporowski A. Use cases for COVID-19 screening and surveillance with rapid antigen-detecting tests: a systematic review. 2021. Manuscript submitted for publication.