Madhukar Pai , César Augusto Ugarte , Alexandra Zimmer , Emily MacLean

Madhukar Pai , César Augusto Ugarte , Alexandra Zimmer , Emily MacLean

IDRC GRANT / SUBVENTION DU CRDI : - VALIDATION OF RAPID MOLECULAR TESTING FOR COVID-19 AND INTEGRATION WITH TUBERCULOSIS DIAGNOSTICS

Project Title

Validation of rapid, molecular testing for COVID-19 and integration with TB diagnostics

Report Type

Final Technical Report

Period Covered

August 1st 2020 – March 31st 2023

Grant Number

Grant # 109554-001

Lead Institution

The Research Institute of the MUHC
2155 Guy Street, Suite 500, Montreal, QC, H3H 2R9, Canada

Collaborating Institution

Universidad Peruana Cayetano Heredia, Lima, Peru

Location of Research

Lima, Peru

Research Team

Dr. Madhukar Pai (PI), McGill University
Dr. César Augusto Ugarte (Co-Investigator), Universidad Peruana Cayetano Heredia
Dr. Nandini Dendukuri (Co-Investigator), McGill University
Dr. Giorgia Sulis (Co-Investigator), McGill University
Dr. Mikashmi Kohli (Co-Investigator), McGill University
Dr. Emily MacLean (Co-Investigator), McGill University
Dr. Charity Oga-Omenka (Co-Investigator), University of Waterloo
Dr. Samira A. Ramini (Co-Investigator), McGill University
Dr. Vania Karami (Co-Investigator), McGill University
Dr. Patricia Espinoza (Co-Investigator), Universidad Peruana Cayetano Heredia
Ms. Alexandra Zimmer (Co-Investigator), McGill University
Ms. Edith Bastidas (Co-Investigator), Universidad Peruana Cayetano Heredia

Submitted: March 30th 2023

Table of Contents

Table of Contents 2
Abstract 3
Keywords 3
The Research Problem 4
1. Validation of rapid, molecular testing for COVID-19 and integration with TB diagnostics 5
1.1 Objectives 5
1.2 Methodology 5
1.3 Project Activities 6
1.4 Project Outputs 7
1.4.1 Deliverables 7
1.4.2 Specific Achievements 8
1.5 Project Outcomes 9
2. Complementary Work to Strengthen Diagnostic Tools for TB and COVID-19 9
2.1 Objectives 9
2.2 Methodology 10
2.3 Project Activities 10
2.4 Project Outputs 11
2.4.1 Deliverables 11
2.4.2 Specific Achievements 12
2.5 Project Outcomes 12
Overall Assessment and Recommendations 13

Abstract

COVID-19 and tuberculosis (TB) are currently the top two infectious diseases by mortality. Universally, personnel and financial resources from existing disease control programs have been diverted to the COVID-19 response. For TB, this has resulted in a steep decline in case notifications, and, for the first time since 2015, annual mortality increased. Strategies to simultaneously address both diseases could help recover missing TB cases and ensure people suffering from either disease receive appropriate care. We have undertaken two initiatives towards integrating testing for COVID-19 and TB in Lima, Peru. First, we investigated rapid testing for COVID-19 and TB in one clinical encounter using a molecular multi-disease testing platform that is available through Peru. Second, we are evaluating the use of cough classification artificial intelligence models, with data collected using a smartphone. Clinical and qualitative aspects of the intervention are being assessed. The findings from our studies will likely be generalizable to other urban settings with high TB burdens.

Keywords

COVID-19; tuberculosis; diagnostics; molecular diagnostics; multiplex testing; artificial intelligence; cough; Peru

The Research Problem

As Peru is dealing with heavy burdens of TB and COVID-19, the country needs an integrated testing solution for both diseases. TB incidence in Peru is approximately 119/100,000, making it a high TB burden country. Since the beginning of the pandemic, there have been over 7 million cases of COVID-19 and approximately 6.6 million deaths attributable to the disease. Data have emerged suggesting that the presence or history of TB may increase the severity of COVID-19 and poor disease outcomes; therefore, testing for both diseases is an important intervention, especially in vulnerable populations such as elderly or immune-suppressed persons.

The GeneXpert platform is available in 16 reference laboratories in Lima and 7 other regions across the country, and Xpert MTB/RIF is recommended for TB diagnosis among key populations (e.g., children, HIV-infected subjects). In countries with such a high TB burden, presumptive COVID-19 cases should also be investigated for TB, given the substantial overlap in clinical presentation and risk factors for developing severe forms. In March 2021, USAID released guidelines for integrated testing for COVID-19 and TB in high TB burden countries. According to USAID, an ideal solution to try to mitigate the burden of TB and COVID-19 would be to test for both diseases in the same clinical encounter using a multiplex system. Similarly, India has introduced a policy recommending bi-directional screening for people with presumptive COVID-19 and TB.

Symptom duration does not rule-out COVID-19 or TB, and there is confusion among healthcare providers in high TB burden settings regarding whether they are dealing with TB or COVID-19. Therefore, aligning efforts to combat both COVID-19 and TB are warranted: diagnostic testing using the GeneXpert system is one such opportunity. In addition, testing for COVID-19 and TB in the same clinical encounter is a patient-centered approach in that it is convenient for the presenting patient, minimizes interactions with the healthcare system, and ideally would reduce the time an infectious person spends in the environment.

As such, the objectives for our primary study are 1) to determine the accuracy of the Xpert Xpress SARS-CoV-2 assay for COVID-19 diagnosis, and 2) to investigate whether a differential diagnosis of COVID-19 or TB can be made with a single sputum sample.

Additionally, we are conducting complementary work investigating the use of novel artificial intelligence (AI) based methods for TB and COVID-19 screening. This complementary study builds upon the emerging field of “acoustic epidemiology” which aims to turn cough sounds into objective biomarkers. As previously mentioned, the challenge of addressing COVID-19 and TB is amplified by the overlapping symptoms, complicating clinical diagnosis. A multi-faceted approach is needed for Peru to get back on track towards TB elimination. Digital cough screening methods could be part of that approach by streamlining patients towards appropriate confirmatory diagnostics.

Determining the success and impact of smartphone-based digital health apps within health systems cannot be limited to monitoring their diagnostic accuracy; we also need to understand the barriers and facilitators to implementation and the acceptability of digital cough monitoring among patients, providers and health systems and surveillance experts. Understanding these perspectives will lead to better, user-centered clinical tools and improved implementation feasibility. We are leveraging the unique opportunity of this prospective study to gain rich, specific, and contextualized insights on the challenges of using already available smartphone apps for digital cough monitoring in a clinical setting and improve implementation approaches.

As such, the objectives of this complementary study are to 1) develop and evaluate the clinical performance of smartphone-based cough classification AI models to screen for TB and COVID-19 and 2)

assess the acceptability and feasibility of using smartphone-based digital cough monitoring technology to screen for TB and COVID-19 among healthcare workers and patients in Lima, Peru.

Since the beginning of this complementary study, additional research has evaluated the use of AI algorithms for TB and COVID-19 screening. One large-scale competition hosted by Sage Bionetworks found that some preliminary algorithms can differentiate TB from non-TB coughs with an accuracy >70% (preliminary results are unpublished). These results support the work being done in this study and will allow me to build upon these preliminary models in an independent cohort of participants.

1. Validation of rapid, molecular testing for COVID-19 and integration with TB diagnostics

1.1 Objectives

Objective 1: To determine the diagnostic accuracy of Xpert Xpress SARS-CoV-2 compared to a reference standard in adults with a clinical picture suggestive of COVID-19 and/or TB at a secondary referral hospital or primary health clinics in Lima, Peru:

- Aim 1.2: To determine Xpress SARS-CoV-2 accuracy using sputum samples

Objective 2: To determine the feasibility of using a single respiratory sample to achieve a differential diagnosis of COVID-19 or TB using Xpert Xpress SARS-CoV-2 and Xpert MTB/RIF, respectively, in adults with a clinical picture suggestive of COVID-19 and/or TB at a secondary referral hospital and primary health clinics in Lima, Peru:

- Aim 2.1: To determine the diagnostic yield on sputum samples of Xpress SARS-CoV-2 and Xpert MTB/RIF Ultra compared to COVID-19 and TB reference standards;
- Aim 2.2.: To obtain qualitative information on the feasibility and acceptability of integrated molecular testing for COVID-19 and TB using the GeneXpert platform

1.2 Methodology

In collaboration with the Instituto de Medicina Tropical Alexander von Humboldt (IMTAvH) at Universidad Peruana Cayetano Heredia (UPCH), we prospectively recruited a cohort of adults with presumptive COVID-19 or TB in Lima, Peru; recruitment sites included a secondary hospital and multiple primary TB clinics. For the primary study and analysis, participants were recruited from February 2021 to April 2022. Participants were included if they had no history of TB treatment in the previous six months and if they reported symptoms suggestive of COVID-19 or TB. This was a modification of the original eligibility criterion for the study. Due to the severe strain on the Peruvian healthcare system, only symptomatic people were admitted to healthcare facilities for much of the pandemic (as asymptomatic persons were considered less at-risk for severe medical complications). Additionally, all participants had to provide a nasopharyngeal (NP) swab and a sputum sample, but performing sputum inductions on asymptomatic participants was not feasible. Together, this means that our findings will be most applicable to settings with symptomatic persons presenting for care. All participants provided written informed consent.

We collected sputum (5 mL) and NP swabs from all participants. NP swabs were placed in approximately 3 mL of transport media. All specimens were transported at 4 degrees to the central study laboratory for processing on the day of collection. The sputum samples were homogenized using glass beads, with 1 mL used for TB testing on Xpert MTB/RIF Ultra (Ultra), a WHO-endorsed, automated PCR assay, and 300

μL used for COVID-19 testing on Xpert Xpress (Xpress) SARS-CoV-2, a FDA emergency-use authorized automated PCR assay run on the same multiplex platform. The remaining sputum was used for smear microscopy and inoculated on BACTEC MGIT liquid culture. 300 μL of NP swab transport medium was also used for COVID-19 testing on Xpress. Laboratory staff were blinded to other participant information.

The sensitivity and specificity and 95% confidence intervals were calculated for Xpress on sputum. As the study progressed, it emerged in the literature that the Xpress test had much lower analytical sensitivity than the PCR test we had planned to use as part of a composite reference standard. The Xpress test on NP swabs has been shown consistently to be one of the most sensitive and specific SARS-CoV-2 detection tests; therefore, we modified our planned analysis for the publication and made Xpress on NP swabs the reference standard for COVID-19. BACTEC MGIT liquid culture served as reference standard for TB.

To assess diagnostic yield of a single sputum sample for the identifying COVID-19 and TB, we compared the proportion of COVID-19 cases identified using sputum on Xpress versus using Xpress on NP swabs. Similarly, for TB, we compared the proportion of TB cases identified using sputum on Ultra versus MGIT liquid culture.

Finally, information regarding integrated molecular testing feasibility and acceptability was obtained through semi-structured interviews with study staff, including members of the microbiology laboratory and study nurses involved in participant recruitment. All interviews were conducted in Spanish.

1.3 Project Activities

In our study population of 600 people with presumed COVID-19 or TB, the prevalence of COVID-19 was 35% and the prevalence of bacteriologically-confirmed TB was 13%. 13 people (2.2%) were concurrently positive on tests for COVID-19 and tests for TB. This suggests that among people with culture-positive TB, COVID-19 prevalence is 14%. This has important ramifications for the clinical management of TB, as steroids that may be used to treatment severe COVID-19 are contraindicated for the treatment of TB. We also described the symptom profiles of people ultimately diagnosed with COVID-19 or with TB. Fatigue, headache, and loss of smell were more common in the former group, but symptoms reported in the latter group typically were longer in duration.

We found that sputum sensitivity on Xpress, compared to Xpress on the manufactured-recommend NP swabs, was 67% (95% CI: 60-73), with specificity of 97% (95% CI: 94-98). Test performance did not vary by age or sex (results not shown). Using a single sputum specimen on the GeneXpert platform, the diagnostic yield of Xpress for COVID-19 was 67%, while the diagnostic yield of Ultra for TB was 96%. As 13/600 participants with bacteriologically-confirmed TB also had a positive result on a test for COVID-19, 46 individuals with presumptive COVID-19 or TB would be tested to find one person with both diseases.

According to interviews with five study staff members, integrated testing was considered highly feasible. The clinical staff at study sites were experienced at collecting samples (both sputum and NP swabs) and had protocols in place for all necessary procedures. This meant that integrated testing did not impose much of a burden on study sites. One challenge identified by study nurses was a reluctance among some people with presumed COVID-19 to undergo TB testing; some patients did not see the relevance of TB to them personally. Laboratory staff reported that integrated testing was very feasible to incorporate into their regular lab practices, although the different GeneXpert test cartridges are highly similar in appearance, leading to some cartridges being wasted.

The investigation of integrated molecular testing was made possible by the funds from IDRC. In particular, funds were utilized to procure a new GeneXpert molecular testing platform, which was used throughout the study's duration for COVID-19 and TB testing. All study participants received a panel of tests to investigate both diseases, which they otherwise would not have received. Additionally, study funds contributed to salary support for clinical and laboratory time taken performing extra work necessary for the study.

We faced a number of modifications and, at times, challenges during the reporting period. There were some notable changes in COVID-19 testing at UPCH that emerged during the pandemic. For example, at the time of grant submission, we had anticipated also using serological tests in our diagnostic work-up of COVID-19. However, throughout 2020 it became known than serological tests are generally unhelpful in diagnosing acute COVID-19 due to their poor diagnostic accuracy, so those tests were never run in our study. At multiple points during the recruitment period, study staff or members of their households tested positive for COVID-19 and had to isolate for two weeks. Obviously, this led to delays in participant recruitment during certain weeks.

To ensure the project's success, holding weekly progress meetings, as well as using informal communication channels (i.e., Whatsapp) throughout the week were important. Additionally, the project succeeded because of the strength of our implementing partners at the Universidad Peruana Cayetano Heredia; the UPCH team was highly experienced and professional. A new international collaboration is challenging in the best of times, but without a strong implementing partner and a committed collaboration, a pandemic-era project may never be realized, regardless of the caliber of project management or idea.

1.4 Project Outputs

1.4.1 Deliverables

Publications

- MacLean EL, Villa-Castillo L, Ruhwald M, Ugarte-Gil C, Pai M. Integrated testing for TB and COVID-19. *Med*. 2022 Mar 11;3(3):162-6.
- MacLean EL, Villa-Castillo L, Espinoza-Lopez P, Caceres T, Sulis G, Kohli M, Pai M, Ugarte-Gil C. Integrating tuberculosis and COVID-19 molecular testing in Lima, Peru: a diagnostic accuracy study. *Lancet Microbe*. Accepted.

Information sharing and dissemination outputs

- Pai M. Calidad en el cuidado de la TB (quality in tuberculosis care). Presentation at: VIII Jornada Científica Internacional Investigación Peruana para el Control de la Tuberculosis. 2021 December 3; Lima, Peru [virtual conference]. This talk was part of the support from McGill to the National TB Program.
- MacLean E. Integration of COVID-19 and tuberculosis testing in Lima, Peru. Presentation at: Infectious Disease and Immunity in Global Health Work-In-Progress Seminar Series. Research Institute of the McGill University Health Centre. 2021 May 19; Montreal, Canada [virtual seminar]. Internal presentation of project and interim results to IDIGH colleagues at the Research Institute of the McGill University Health Centre.
- MacLean E. Integration of COVID-19 and tuberculosis testing in Lima, Peru. Presentation at: Infectious Disease and Immunity in Global Health Trainee Research Day. Research Institute of the

McGill University Health Centre. 2021 May 21; Montreal, Canada [virtual research day]. Internal presentation of project and interim results to trainee colleagues and faculty at annual research day event.

- Ugarte-Gil C and MacLean E. Integration of COVID-19 and tuberculosis testing in Lima, Peru. Presentation at: Integrating COVID-19 and TB Testing. McGill Summer Institute in Infectious Disease and Global Health. 2021 June 11; Montreal, Canada [virtual course]. International lecture of progress in integrating COVID-19 and TB testing, as well as project and interim results at McGill Summer Institute, a two-week long symposium of international researchers, ministry of health representatives, patient advocates, industry personnel, and graduate students.
- MacLean E, Villa-Castillo L, Cáceres T, Ugarte-Gil C, Pai M. Integration of COVID-19 and tuberculosis testing in Lima, Peru. Presentation at: 52nd Union World Conference on Lung Health. 2021 October 19-22 [virtual conference].
- Ugarte-Gil C. Integration of COVID-19 and tuberculosis testing in Lima, Peru. Presentation at: Jornada Cientifica: Investigación Peruana para el control de la tuberculosis. 2021; Lima, Peru [virtual conference]. Presentation of project and results at Peruvian annual national conference on tuberculosis.
- MacLean, E and Ugarte-Gil C. Integrating COVID-19 & TB Testing (Update). Presentation at: McGill Summer Institute in Infectious Diseases and Global Health. 2022 June 10; Montreal, Canada [virtual course]. International lecture of final results of integrating COVID-19 and TB testing at McGill Summer Institute, a two-week long symposium of international researchers, ministry of health representatives, patient advocates, industry personnel, and graduate students

1.4.2 Specific Achievements

Completed training, capacity-building, institutional reinforcement, and increased research skills impact

- Procurement and installation of GeneXpert System (Cepheid, USA)
  - This automated PCR module has been incorporated into the laboratory workflow at the Instituto de Medicina Tropical Alexander von Humboldt (IMTAvH). The IMTAvH provides support to the public hospital where it is located (Hospital Nacional Cayetano Heredia) and other public institutions. The addition of this machine has allowed the laboratory to increase their GeneXpert-based testing capacity by 33%, and this state-of-the-art module will be compatible with the upcoming TB tests for drug resistance and host biosignatures.
- Training and network strengthening with Cepheid
  - Virtual meeting and training session with local representative from Cepheid (manufacturer of GeneXpert) with lab personnel at IMTAvH laboratory were held. As well as sharpening skills, laboratory personnel re-established direct connections with manufacturer and support staff.
- Increased research or administrative skills
  - Skills acquired and sharpened by various team members include grant, report, and IRB application writing skills; remote study management and collaboration; development of protocols and quality assurance plans; utilization of and survey design in REDCap; data input and quality control; collecting follow-up information;
  - Training in COVID-19 biosafety measures, molecular testing for COVID-19 with GeneXpert.
- Graduate trainees

- This project was part of a PhD trainee's (Emily MacLean) thesis. Dr. MacLean completed her PhD at McGill University under the supervision of Dr. Madhukar Pai in March 2022, and her manuscript related to this project has been accepted in *Lancet Microbe* (see section 1.4.1).
- Luz Villa-Castillo, a core member of the UPCH team during this project, will join McGill University in September 2023 to complete an MSc in epidemiology under the supervision of Dr. Madhukar Pai.

1.5 Project Outcomes

Our study is the largest, and one of the first, to evaluate the performance of Xpert Xpress SARS-CoV-2 (Xpress) in sputum. Understanding the performance of Xpress SARS-CoV-2 in sputum is important given continued sporadic supply-chain disruptions and medical equipment stock-outs; in case of NP swab unavailability, there would be great utility in using alternate specimens for testing, such as sputum. We expect that our findings regarding test accuracy are generalizable to other urban high TB burden settings where symptomatic individuals present for care.

Our work contributes to the growing body of literature around TB/COVID-19 interactions and serves as key evidence for countries in deciding whether to introduce or implement strategies of bi-directional testing for COVID-19 and TB. This is particularly timely, given that influential stakeholders such as Global Fund and USAID have published policy briefs recommending integrated testing for COVID-19 and TB. However, aside from our (soon to be published) study, evidence to support this policy are scarce. Although the published policies for integrated testing are meant to find the most vulnerable people missing out on proper healthcare, our work indicates that uniformly testing all people with presumed COVID-19 or TB for both diseases is not the ideal approach, and a more finely tuned policy is needed to maximize the strategy's potential. Our study adds important nuance to the discussion around implementing integrated testing.

This is critical information for National TB Programs (NTP) with respect to planning and resource allocation, particularly as many NTPs already have existing GeneXpert networks, but NTP budgets are limited. Integrated testing may be very feasible, but in order to have the largest possible impact (in terms of finding people with COVID-19 and/or TB), it needs to be thoughtfully implemented – perhaps during COVID-19 surges, as opposed to constantly. This underscores the need for strong surveillance systems, which could inform when integrated testing should be implemented.

Finally, we are continuing to conduct follow-up phone calls with all of our study participants. The results obtained from these calls will provide some insight into the post-study health status of those who tested positive for TB and concurrent COVID-19 and TB, which will be high interest to NTPs in high TB burden countries. This may be particularly relevant to IDRC, as the financial burden, in addition to the physical and mental toll, imposed by TB on patients is severe.

2. Complementary Work to Strengthen Diagnostic Tools for TB and COVID-19

2.1 Objectives

Objective 1: To develop and evaluate the clinical performance of smartphone-based cough classification AI models to screen for, and hence differentiate, TB and COVID-19

Objective 2: To involve end-users in assessing the feasibility and acceptability of using smartphone-based digital cough monitoring screening of TB and COVID-19 in Lima, Peru

2.2 Methodology

In collaboration with the Universidad Peruana Cayetano Heredia, we prospectively recruited a cohort of 150 adults presenting with a cough at primary health clinics in Lima, Peru (as part of the Xpert validation study presented previously). Using a smartphone with the Hyfe Research application, participants recorded 5-10 solicited coughs upon enrollment. Patients received reference standard testing for both respiratory diseases. For COVID-19, a composite reference standard was performed, including clinical symptoms, epidemiological history, and GeneXpert SARS-CoV-2 PCR. For TB, we used MGIT liquid culture. Participants were then given the option to carry a smartphone over the next 7 days to passively record spontaneous (unsolicited) coughs. This was not a requirement to participate in the study and only a subset of patients agreed to participate in this follow-up. Recording strategies were designed to optimize recording while preserving privacy. Reference standard testing and patient sociodemographic/ clinical information were obtained from the validation parent study. A cough database was generated. Annotations with diagnosis, demographic, and clinical data is being finalized. Using this database, we will develop AI models using deep learning methods by applying a standard random split-sample approach (90% derivation, 10% validation) to train TB and COVID-19 cough classifiers and assess models' performance (sensitivity, specificity, receiver operating curve).

Success of such an innovative approach was also contingent on how patients and providers engaged with technology and perceived its impact. This work thus included a nested qualitative study, involving 6 in-depth interviews with purposively selected healthcare staff. We administered a short "Acceptability Questionnaire" to participants that recorded cough sounds at the facility and who participated in the 7-day follow-up. Patient focus group discussions (FGDs) are being finalized over the next two months. These discussions will be held by FGDs will be conducted separately by gender (male, female) and stratified by age group (<35 and ≥35), with each FGD containing 4-8 individuals. We will also conduct the FGDs within different patient groups: 1) refused to participate in the cough sub-study but participated in the parent study; 2) only agreed to cough 5-10 times at the facility; and 3) agreed to participate in the 7-day follow-up study. In the event that we are unable to obtain 4-8 participants, we may merge participants across gender and/or age groups.

Audio recordings from the FGDs and interviews will be transcribed verbatim, checked for accuracy, anonymized, and thematically analyzed using a grounded theory framework, to discern recurring patterns and themes related to the acceptability of cough recording technology, stigma, data privacy concerns, and the clinical utility of cough monitoring tools. Data will be organized, coded, and categorized iteratively using Quirkos software by two study researchers who will practice critical reflexivity and crosscheck emerging codes in consultation with the study team to enhance inter-rater reliability and ensure that the analysis was grounded in participants' narratives. Codes will be further refined and contextualized through a process of constant and discursive comparative analysis across interviews and FGD transcripts (triangulation) to facilitate qualitative thematic analysis.

2.3 Project Activities

Recruitment of the 150 participants concluded in January 2023. Of the 150 participants, 30 individuals were involved in the longitudinal 7-day follow-up. Clinical data entry has been setback due to the unforeseen political crisis that occurred in December 2022. On the AI side, our team is currently working

on extracting cough features and building the cough classification algorithm is ongoing, therefore no final model is available at the moment. We anticipate having final models ready by end of 2023.

We conducted in-depth interviews with six healthcare workers. The transcripts of these interviews are currently being translated from Spanish to English. Preliminary findings from these interviews suggest that there is a good degree of acceptability of the technology on behalf of the healthcare workers that used the smartphones to enroll patients. Observations from the healthcare workers regarding patient acceptability were also generally positive. Some concerns were raised for some patients who lacked trust in the technology or found that coughing was painful or irritating and therefore did not wish to participate in the complementary study. Finally, the healthcare workers observed that most patients opted out of the 7-day longitudinal follow-up because they were uncomfortable carrying a smartphone with them, raising concerns about safety and fear or breaking or losing the smartphone. This may inform future studies to adapt the recording modality to something less obstructive and obvious. Of note, no smartphones were stolen, lost, or damaged during this study.

With respect to feasibility, the healthcare workers noted that some patients, particularly elderly patients, did not feel comfortable operating a smartphone and therefore declined to participate in the 7-day follow-up. Smartphone charging and cell reception were not a feasibility concern in this population.

Additional information regarding patient feasibility and acceptability will be obtained from the Acceptability Questionnaire (data entry still ongoing) and from patient focus group discussions. The FGDs will be conducted between mid-March until end of April, and will involve different groups: 1) patients who only coughed 5-10 times at the facility, 2) patients who participated in the 7-day follow-up and 3) patients who refused to participate in the sub-study.

2.4 Project Outputs

2.4.1 Deliverables

Publications

- Planned manuscript detailing the AI algorithms developed and evaluating their diagnostic accuracy for screening TB and COVID 19.
- Planned manuscript on the qualitative data collected from healthcare workers and patients on the acceptability and feasibility of using this technology in this population.
- Zimmer AJ, Ugarte-Gil C, Pathri R, Dewan P, Jaganath D, Cattamanchi A, Pai M, Grandjean Lapierre S. Making cough count in tuberculosis care. Comms Med. 2022 2(1);1-8. [note: although not an output of this project, this paper from our team is relevant to our complementary work to strengthen diagnostic tools for TB and COVID-19]

Information sharing and dissemination outputs

- Zimmer A. Smartphone-based artificial intelligence cough classification for tuberculosis and COVID-19 screening in Lima, Peru. Poster. McGill Global Health Night 2022 November 9; Montreal, Canada.
- The collection of these cough sounds along with the extensive clinical metadata will allow us to build a large, open-source database of labelled cough data that can be used by researchers to train and validate future AI models.

2.4.2 Specific Achievements

Completed training, capacity-building, institutional reinforcement, and increased research skills impact

- March 2022: Training with Hyfe on how to install and operate the Hyfe Research application
    - This was a virtual training session with the Hyfe team, UPCH project manager (Dr Patricia Espinoza), and McGill project manager (Ms. Alexandra Zimmer). Hyfe walked us through the steps on how to install the Hyfe Research app as well as configure the smartphone into a "locked" mode so that the participants could not use the smartphone for other purposes.
- March-April 2022: Training of data collection research nurses
    - UPCH project manager (Dr Patricia Espinoza) coordinated a training session with the six research nurses who were involved in collecting the cough data. All six research nurses were shown how to operate the smartphone, use the Hyfe Research app, and record cough sounds. They were also trained on how to train patients how to use the smartphone for the 7-day follow-up.
- November 2022: Qualitative training on conducting in-depth interviews
    - Local qualitative researcher (Ms. Edith Bastidas), UPCH project manager (Dr Patricia Espinoza), and McGill project manager (Ms. Alexandra Zimmer) conducted an in-person training on how to lead qualitative in-depth interviews for a group of local research nurses.
- January 2023: Qualitative training on conducting focus group discussions
    - Local qualitative researcher Ms. Edith Bastidas), UPCH project manager (Dr Patricia Espinoza), and McGill project manager (Ms. Alexandra Zimmer) conducted an in-person training on how to lead qualitative patient focus group discussions for a group of local research nurses.
- Graduate trainees
    - This project is part of a PhD trainee's (Alexandra Zimmer) thesis. Ms. Zimmer is expected to complete her PhD in 2024 at McGill University under the supervision of Dr. Madhukar Pai.

2.5 Project Outcomes

For the complementary cough sub-study, we introduced a new way of using smartphone technology to collect data from patients. Based on narrative reports from healthcare workers, this kind of healthcare technology is not used within these low-resource settings. However, the integration of novel technologies into routine healthcare practice will grow as technology continues to advance. Therefore, this early introduction to collecting cough sounds using a smartphone application may help promote future practices involving the collection of personal health data and help influence behaviors regarding smartphone technology as a biomedical device.

The final output of this research — AI algorithms for screening COVID-19 and TB — will be preliminary. A sample size of 150 participants is too small to develop a final AI model that can be deployed. However, the capacity building, resource acquisition, and training that this study invested in study has allowed the local research team at UPCH to apply for additional funding opportunities to expand recruitment. These future studies in this setting will also expand the study to investigate longer follow-up times among TB and COVID-19 patients to using cough frequency as a passive biomarker for monitoring treatment response and disease progression.

The qualitative component of our study sets it apart from other AI cough studies, as well as other diagnostic accuracy studies. Too often there is a lack of context regarding acceptability and feasibility of integrating novel diagnostics or technology into a healthcare setting. Having health workers and patients provide their feedback on the technology early in the development process is essential feedback that will help shape a person-centered product and inform implementation strategies.

For this complementary work, it is important to recognize that biological sex affects respiratory physiology and voice tone, which may impact cough sounds and patterns. Imbalanced recruitment of biological males and females in the study population risks creating AI cough models. To ensure female representation in our sample and prevent AI gender bias, we incorporated sex/gender variables in the design, data collection, and analytic stages of our study. For AI model development, we will re-weigh our samples to ensure equal representation of males and females, preventing biased algorithms. We will also perform subgroup analyses within each sex to evaluate the extent to which acoustic signatures differ and how well the algorithm performs for each sex.

We also qualitatively captured gender-specific concerns related to the technology through the in-depth interviews and focus group discussions, allowing us to understand factors that may influence acceptability and feasibility of using this technology across sex and gender. In doing so, we aim to create a user-centered tool that can help bridge the digital gender divide.

We are currently applying for a follow-up grant, funded by the Ai4PEP Network “Household screening for contagious and transmissible respiratory infections using artificial intelligence-based cough monitors” with our partners at UPCH and the Centre Hospitalier de l’Université de Montréal.

Overall Assessment and Recommendations

Through the projects funded by this IDRC grant, the McGill and UPCH teams have established a strong collaboration that will lead to other research projects, knowledge sharing, and personnel exchanges in the future. The results of the primary study have been shared with the Peruvian Ministry of Health to assist in their policy development. The work has also been useful as it has contributed to the scarce evidence base around integrated COVID-19 and TB testing; novel approaches for finding TB are particularly critical now, given that many people who developed TB during the height of the COVID-19 pandemic never obtained a diagnosis.

Although perhaps obvious in retrospect, for future projects, lengthening the timelines to allow for pandemic-related disruptions to the study (e.g., staff becoming ill with COVID-19 and being unable to work) would make sense. Delays related to importing test cartridges also slowed down initiating the study; more lead time should be scheduled to allow for such logistical issues in the future.

We were able to investigate multiple topical questions in the field of molecular and next-generation infectious disease diagnostics, i.e., integrating COVID-19 and TB testing through multiplex platform testing, as well as through an AI-aided cough detection app. Inventive strategies like these are intuitively appealing and make sense; as such, major funders and other stakeholders may want to quickly implement them without much examination Interestingly, our work on integrating COVID-19 and TB molecular testing shows that is not a perfect approach, with TB testing being less acceptable than testing for COVID-19 to patients. The qualitative work incorporated into the AI cough detection study will help derive a deeper understanding of patient views of this technology. In the future, when assessing a new diagnostic technology, obtaining qualitative inputs from patients, alongside quantitative

performance characteristics, would be essential. Overall, we believe our work was valuable, with efforts commensurate to knowledge generated.

Future work should consider integration of not only COVID-19 and TB, but also other circulating respiratory pathogens in Peru. This type of work is lacking within the acoustic epidemiology field, where most studies have investigated the etiology of cough for one disease (usually COVID-19). However, many diseases may cause cough. Efforts to address two key pathogens in this study (TB and COVID-19) improve upon prior research, however future work should incorporate reference standard testing for other diseases such as influenza and RSV.

Conducting collaborative global health research amidst a global pandemic has been challenging, but also lessons have been learned. The IDRC team were very clear in communicating with us what their needs and expectations were. The flexibility and open-mindedness of IDRC, e.g., permitting no-cost extensions and updates to the project as the field evolved, is greatly appreciated and should be encouraged.