Check for updates

STUDY PROTOCOL

REVISED Design and protocol for a pragmatic randomised study to optimise screening, prevention and care for tuberculosis and HIV in Malawi (PROSPECT Study) [version 3; peer review: 2 approved]

Peter MacPherson [1,2], Emily L. Webb [3], David G. Lalloo [1], Marriott Nliwasa[2,4], Hendramoorthy Maheswaran [5], Elizabeth Joekes[1], Dama Phiri [6], Bertie Squire [2], Madhukar Pai[7], Elizabeth L Corbett [3]

[1]Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, UK
[2]Malawi-Liverpool-Wellcome Trust Clinical Research Programme, Blantyre , Malawi
[3]London School of Hygiene & Tropical Medicine, London, UK
[4]Helse Nord TB Programme, Department of Microbiology, College of Medicine, University of Malawi, Blantyre, Malawi
[5]Department of Public Health and Policy, University of Liverpool, Liverpool, UK
[6]Radiology Department, Queen Elizabeth Central Hospital, Blantyre, Malawi
[7]McGill International TB Centre, McGill University, Montreal , Canada

Abstract

Background: Adults seeking diagnosis and treatment for tuberculosis (TB) and HIV in low-resource settings face considerable barriers and have high pre-treatment mortality. Efforts to improve access to prompt TB treatment have been hampered by limitations in TB diagnostics, with considerable uncertainty about how available and new tests can best be implemented.
Design and methods: The PROSPECT Study is an open, three-arm pragmatic randomised study that will investigate the effectiveness and cost-effectiveness of optimised HIV and TB diagnosis and linkage to care interventions in reducing time to TB diagnosis and prevalence of undiagnosed TB and HIV in primary care in Blantyre, Malawi. Participants ($\geq$ 18 years) attending a primary care clinic with TB symptoms (cough of any duration) will be randomly allocated to one of three groups: (i) standard of care; (ii) optimised HIV diagnosis and linkage; or (iii) optimised HIV and TB diagnosis and linkage. We will test two hypotheses: firstly, whether prompt linkage to HIV care should be prioritised for adults with TB symptoms; and secondly, whether an optimised TB triage testing algorithm comprised of digital chest x-ray evaluated by computer-aided diagnosis software and sputum

Open Peer Review

Approval Status ✔ ✔

	1	2
version 3 (revision) 21 Nov 2018		✔ view
version 2 (revision) 30 Jul 2018	✔ view	
version 1 18 May 2018	? view	? view

1. **Frank Cobelens**, Academic Medical Center, Amsterdam, The Netherlands
   Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands

2. **Tom A. Yates** [iD], Imperial College London,

GeneXpert MTB/Rif can outperform clinician-directed TB screening. The primary trial outcome will be time to TB treatment initiation by day 56, and secondary outcomes will include prevalence of undiagnosed TB and HIV, mortality, quality of life, and cost-effectiveness.

Conclusions: The PROSPECT Study will provide urgently-needed evidence under "real-life" conditions to inform clinicians and policy makers on how best to improve TB/HIV diagnosis and treatment in Africa.

Clinical trial registration: NCT03519425 (08/05/2018)

Keywords
Tuberculosis, HIV, public health, screening, diagnostics, treatment, randomised controlled trials, sub-Saharan Africa

London, UK

Any reports and responses or comments on the article can be found at the end of the article.

This article is included in the Malawi-Liverpool Wellcome Trust Clinical Research Programme gateway.

Corresponding author: Peter MacPherson (peter.macpherson@lstmed.ac.uk)

How to cite this article: MacPherson P, Webb EL, Lalloo DG *et al*. **Design and protocol for a pragmatic randomised study to optimise screening, prevention and care for tuberculosis and HIV in Malawi (PROSPECT Study) [version 3; peer review: 2 approved]** Wellcome Open Research 2018, **3**:61 https://doi.org/10.12688/wellcomeopenres.14598.3

First published: 18 May 2018, **3**:61 https://doi.org/10.12688/wellcomeopenres.14598.1

Introduction

Tuberculosis (TB) is now the leading infectious cause of death worldwide[1]. In 2016, there were an estimated 1.4 million deaths attributed to tuberculosis global, with an additional 0.4 million deaths from TB among people living with Human Immunodeficiency Virus (HIV) infection[1,2]. The countries of sub-Saharan Africa have been disproportionately affected by the HIV-TB co-epidemics. Following extremely rapid increases in TB incidence, prevalence and deaths during the 1990s and 2000s in the region that occurred concurrently with rapid increases in population HIV prevalence[3], TB rates have only begun to decline in the region in recent years[1]. Although the expansion of coverage of effective antiretroviral therapy (ART) for treatment of HIV in many sub-Saharan countries has likely contributed to recent reductions in mortality, the pace of decline is unacceptably slow.

New impetus has been given to efforts to improve tuberculosis control by the recent-agreed global End-TB Strategy[4]. This strategy, which was endorsed by WHO in 2015, demands global action and intensified research to address HIV-associated TB in 30-high HIV/TB burden countries that together comprise 87% of the global burden of TB[2]. Key targets for the End-TB strategy include achievement by 2035 of a 90% reduction in TB incidence and a 95% reduction in TB mortality compared to 2015[4].

Modelling studies have shown however that the End-TB targets will not be met without a step-change in efforts to improve the early diagnosis and effective treatment of all individuals with TB[5]. Of concern remains low population TB case detection rates, and high case-fatality ratios, particularly among people living with HIV[6].

Adults seeking care at health facilities in sub-Saharan Africa are an important group to address in TB care and prevention programmes, as they have high prevalence of undiagnosed TB[7], a substantial burden of undiagnosed and untreated HIV[8], and high mortality rates if not promptly diagnosed and linked to treatment[9].

Our previous studies - similar to research from other countries in sub-Saharan Africa - have shown that the patient pathway from first health centre attendance, through diagnosis to successful treatment outcome is tortuous, with high rates of drop-out from care[8,10–12]. Importantly, as well as having high mortality rates, individuals with symptoms of pulmonary TB who are not rapidly diagnosed may continue to transmit TB to others in the community, further limiting control efforts.

Although WHO guidelines promote intensified case-finding for TB among adults attending health facilities in high TB-prevalence settings[13], implementation of routine screening for TB is known to be suboptimal in many settings. Exit interviews done with patients leaving health facilities have shown that clinicians rarely conduct an initial symptom screen[14,15]. Moreover, even when symptoms of presumptive TB are reported, only a small fraction receive appropriate investigations for TB[14,15]. Thus, even at the earliest step of the TB diagnostic and care pathway, there are high rates of loss from the cascade.

In addition to low rates of TB screening, we have previously shown that only a small proportion (13%) of adults attending health facilities receive HIV testing, despite WHO and Malawi guidelines recommending a strategy of universal provider-initiated HIV testing and counselling (HTC) for all individuals attending health centres, regardless of reason[8]. Uptake of HTC was highest among pregnant women attending for routine antenatal care, where considerable efforts have been undertaken to operationalise universal HTC as part of prevention of mother to child HIV transmission programmes. However, other groups, such as men and non-pregnant women have considerably lower rates of HIV testing completion[8]; as they are attending with an acute care episode and have a higher prevalence of active tuberculosis, they may have substantially worse outcomes compared to pregnant women.

Through systematic reviews, meta-analysis, and prospective cohort studies we have shown that even if patients attending health centres are diagnosed with TB or HIV, they face considerable barriers to treatment initiation. In Africa, 18% of adults with sputum smear-positive tuberculosis will not initiate tuberculosis treatment promptly[12], whilst only a fifth of HIV-diagnosed adults will remain in care continuously to initiation of ART[10]. Across both conditions a number of common factors hindering access to treatment have been identified, including: requirements to make multiple health centre visits for registration and assessment visits; debility; competing demands, including from work and education; and high out-of-pocket costs associated with visiting health centres[16].

When both HIV and TB are suspected or diagnosed, patients can face even greater challenges. Despite repeated calls for integration of HIV and TB care and prevention services, most clinic services remain vertically-organised. This means that patients often require multiple health centre visits on different days of the week to receive HIV and TB assessments and treatment, multiplying their adverse case-seeking costs and potentially worsening outcomes[6,16].

We, and other, have therefore argued that new approaches are required to improve integration of HIV and TB screening, prevention and care services in health facilities in Africa that can provide same-day, same-clinic diagnosis and treatment linkage for both conditions at minimum inconvenience to patients[16]. Such an approach, if effective, is likely to have large benefits for patients, public health, and for health systems by improving case detection and treatment access, reducing mortality, and mitigating the catastrophic costs associated with care-seeking. However, strong evidence for effectiveness

obtained through robust randomised controlled trials is currently lacking.

Current TB screening approaches are reliant upon diagnostic tests with considerable limitations.

Sputum smear microscopy has been the mainstay of investigations for pulmonary tuberculosis for nearly a century. Although specificity is high, sensitivity remains unacceptably low, especially among HIV-positive adults[17,18]. Moreover, as sputum smears must be prepared, fixed and examined under light or fluorescence microscopy, infection control, quality control, throughput, and achievement of same-day diagnosis are challenging.

TB culture of sputum is slow (3–8 weeks, even with automated liquid culture systems), and relies upon availability of a high-quality laboratory. These requirements mean that whilst culture has importance for individual management of complex cases and for monitoring and evaluation, it is not practical as a point of care test[19].

GeneXpert MTB/Rif is an integrated and automated cartridge-based nucleic acid amplification test that can provide results for the detection of *M tuberculosis* and associated rifampicin resistance within two hours[20]. There are two components to the test: the cartridge in which the biological sample is added to the assay, and a standalone unit in which cartridge is placed and where the nucleic acid amplification and detection takes place. The sensitivity of the Xpert assay is substantially higher than sputum smear microscopy[20], and the newest version (Xpert Ultra) shows pooled sensitivity (among HIV-positive and HIV negative samples) that is 5% higher than the first-generation assay, with a 12% gain among HIV-positive adults. WHO has endorsed GeneXpert MTB/Rif as the first line test among adults suspected to have multidrug resistant TB or HIV-associated TB.

Despite these advantages, there are some barriers to the widespread implementation of Xpert in low-resource, high TB prevalence settings. In particular, even at current concessional pricing for low-resource settings, Xpert is prohibitively expensive as a first line test for most national programmes.

Chest radiography has high sensitivity for pulmonary TB even in HIV coinfection[21,22], and continues to play an important role in TB diagnosis in high-income settings. Although chest x-ray has been used for many years as a diagnostic tool (usually at the end of screening algorithms), widespread implementation in high prevalence settings has been limited by poor access to high quality equipment and expert radiologists, low specificity (leading to over-diagnosis of TB if chest x-ray alone is used) and high inter-reader variability[22]. Recent advances in digital chest x-ray technologies have reinvigorated interest in the use of chest x-ray as an initial triage tool in primary care in Africa.

Chest X-ray may also be used as a triage test for TB[22]. In this triage approach, individuals with any abnormality identified on chest x-ray undergo confirmatory microbiological testing. Using a point-of-care high specificity molecular sputum testing for confirmatory testing (e.g. Xpert MTB/Rif) could allow accurate same-day TB diagnosis and treatment initiation in primary care.

In December 2016, WHO released a new evidence review and guidance[22] for chest x-ray TB triage that used data from systematic review to model the potential effectiveness of TB screening algorithms, and showed that triage using chest x-ray, followed by GeneXpert MTB/Rif could substantially outperform other approaches.

Currently, countries such as Malawi have low coverage of radiology services, including trained radiologists. Computer-assisted detection (CAD) software - statistical algorithms used to classify digital images - is now available, and can be integrated within new digital x-ray units to provide immediate triage[23]. WHO recently systematically reviewed available evidence for one CAD system (CAD4TB, Delft Imaging Systems, Netherlands). Across 13 studies conducted in a variety of populations, sensitivity was as high as reading by radiologists, although specificity was lower necessitating microbiological confirmation of TB[22]. Whilst promising, WHO recommends that *"CAD can be used for TB detection for research, ideally following a protocol that contributes to the required evidence base for guideline development"*[22].

In summary, adults with symptoms of tuberculosis in Malawi face considerable health systems delays, large out-of-pocket expenses, and have a high risk of mortality before diagnosis and treatment. To achieve the End-TB Strategy goals, a package of same-day, same-clinic diagnosis and treatment linkage interventions for both TB and HIV are urgently required. In an individually-randomised, open, three-arm controlled trial, The PROSPECT Study will investigate whether optimised TB and HIV diagnosis and treatment linkage interventions are cost-effective in reducing time to TB treatment initiation, and in improving case detection.

Methods
Study design
The PROPSECT Study is an open, three-arm pragmatic randomised controlled trial.

Study hypothesis
The PROSPECT Study will test the hypothesis that an optimised same-day TB/HIV screening and treatment linkage intervention for adults with presumptive tuberculosis in primary care could result in important improvements in eight-week case detection, treatment initiation and mortality.

Objectives
I. Among adults with TB symptoms attending primary care in Malawi, to investigate the effectiveness of an optimised same-day screening algorithm consisting of rapid HIV testing, computer-assisted CAD4TB chest x-ray triage and, if abnormal, Xpert MTB/Rif rapid sputum molecular testing, and linkage to treatment.

II. In a nested diagnostic accuracy study evaluate the sensitivity and specificity of computer-assisted chest x-ray triage compared to classification by radiologists and bacteriological diagnosis.

III. Undertake a cost-utility analysis of the PROSPECT interventions to estimate the incremental cost per QALY gained from providing optimised TB and HIV diagnosis and linkage to care.

Study site and population

The study will be done at Bangwe Health Centre, a busy health centre located in a densely populated urban neighbourhood of Blantyre, Malawi. Adult prevalence of tuberculosis and HIV in Blantyre are >900 per 100,000[24] and 18.5% respectively[25].

TB diagnostics available at the study clinic include sputum smear fluorescence microscopy with LED microscopes, and Xpert MTB/Rif. Although Malawi and World Health Organization (WHO) guidelines recommend screening for tuberculosis for all individuals attending health facilities with acute illness (with Xpert preferred as the first-line investigation for symptomatic HIV-positive individuals), our previous research shows that TB diagnostics are underutilised[26].

At the study clinic, comprehensive HIV care is available and includes: routine provider-initiated HIV testing and counselling, screening and treatment of opportunistic infections, provision of chemoprophylaxis, and treatment with antiretroviral therapy. Malawi National Guidelines have recommended a "test and treat" approach to HIV since 2015, with all individuals diagnosed with HIV being eligible for antiretroviral therapy. Our previous research has shown however that coverage of HIV testing and rates of linkage to antiretroviral therapy are suboptimal[8,10], as in many other African settings.

Study participants will be adults aged 18 years or older who attend Bangwe Health Centre for acute care with symptoms of tuberculosis (cough of any duration). As this is a pragmatic randomised trial that aims to provide evidence for policymakers under "real-life" conditions, eligibility criteria will be broad, and will reflect the characteristics of adults attending primary health centres in Southern Africa to maximise generalisability. We will exclude: individuals who are taking treatment for tuberculosis, or who have taken tuberculosis treatment in the preceding 6-months; individuals taking isoniazid preventive therapy; and individuals who live outside of Blantyre or plan to relocate outside of Blantyre in the next six months.

Research assistants based at the clinic registration desk will screen all daily attenders against eligibility criteria. As individuals may attend the clinic on more than one occasion during the study period, the research assistant will record a digital fingerprint (Simprints, Cambridge, UK) from all clinic attenders to ensure that repeat clinic attendance episodes are recorded and so removing the potential for duplication in trial recruitment. Where the number of eligible clinic attenders exceeds recruitment capacity, we will recruit participants up to a daily limit, with details finalized pending completion of pilot work showing the number of eligible participants per day."

Randomisation and blinding

Participants will be individually randomly allocated in a 1:1:1 ratio into one of three groups.

Enrolment and baseline questionnaires will precede randomization. Randomisation will be done by research assistants using a random number allocation schedule running on study data-collection electronic tablets. Because of the nature of the study and the interventions offered, it will not be possible to blind participants or research assistants to allocation groups. Nevertheless, extensive steps will be taken to ensure that research assistants undertaking outcome assessments are blinded to participants' group. Additionally, the investigators, including the chief investigator and trial statistician, will remain blinded to allocation groups until final analysis. No unblinded interim analysis will be conducted.

Interventions

All participants

All participants will complete a baseline questionnaire (Supplementary File 1), that will record demographic and clinical characteristics (including previous HIV and TB care), as well as geolocation information[16] to facilitate home tracing if participants don't attend for outcomes assessments. All participants will additionally complete the EuroQoL EQ5D (Chichewa) tool (English version validated for the UK available here) to measure health-related quality of life.

Group 1: Standard of care

Interventions available to participants allocated to Group 1 are intended to mirror the current standard of care for HIV and TB screening and linkage to care under routine conditions in primary care in Malawi (Figure 1). This will ensure that the incremental cost-effectiveness of interventions offered in Groups 2 and 3 can be compared to a standard care that reflects "real-life" practice.

Participants allocated to Group 1 will be directed to the routine facility waiting area to be seen by facility health workers. Screening for HIV and tuberculosis will be directed by the routine facility health workers in accordance with Malawi National Guidelines, and without any further trial interventions.

Group 2: Optimised HIV screening and linkage to care

Participants allocated to Group 2 will be offered a supervised HIV self-testing intervention using oral fluid rapid diagnostic kits (OraQuick® HIV-1/2 rapid antibody test kits, manufactured in Thailand for: OraSure Technologies, Inc. Bethlehem, PA, USA). Participants will have their identity validated by fingerprint scanning, will be provided with brief pre-test counselling and instruction and demonstration in the use of the oral fluid kits by research assistants, and will be asked to self-test in a private clinic area, with support from the research assistant.
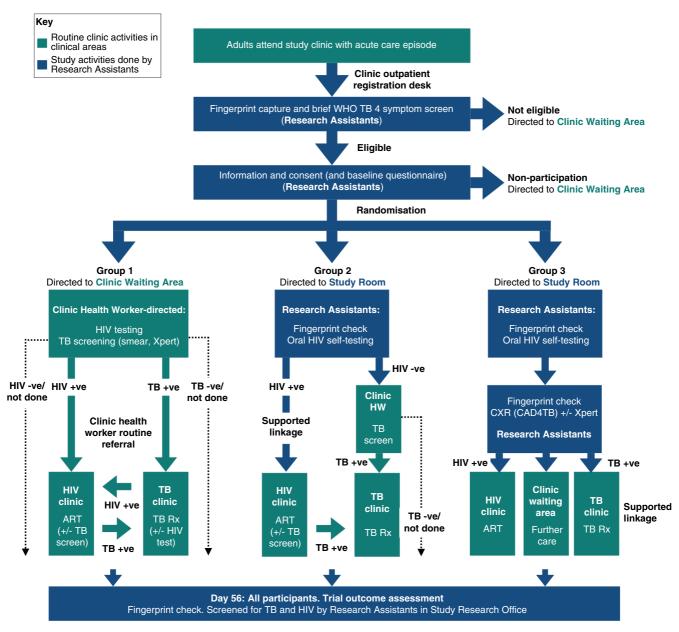
Key
- Routine clinic activities in clinical areas
- Study activities done by Research Assistants

Adults attend study clinic with acute care episode

Clinic outpatient registration desk

Fingerprint capture and brief WHO TB 4 symptom screen **(Research Assistants)**

Not eligible Directed to Clinic Waiting Area

Eligible

Information and consent (and baseline questionnaire) **(Research Assistants)**

Non-participation Directed to Clinic Waiting Area

Randomisation

Group 1 Directed to Clinic Waiting Area

Clinic Health Worker-directed:
HIV testing
TB screening (smear, Xpert)

HIV -ve/ not done | HIV +ve | TB +ve | TB -ve/ not done

Clinic health worker routine referral

HIV clinic ART (+/- TB screen) — HIV +ve — **TB clinic** TB Rx (+/- HIV test)

TB +ve

Group 2 Directed to Study Room

Research Assistants:
Fingerprint check
Oral HIV self-testing

HIV +ve | HIV -ve

Supported linkage

Clinic HW TB screen

TB +ve | TB -ve/ not done

HIV clinic ART (+/- TB screen) — TB +ve — **TB clinic** TB Rx

Group 3 Directed to Study Room

Research Assistants:
Fingerprint check
Oral HIV self-testing

Fingerprint check
CXR (CAD4TB) +/- Xpert
Research Assistants

HIV +ve | | | TB +ve

HIV clinic ART | **Clinic waiting area** Further care | **TB clinic** TB Rx

Supported linkage

Day 56: All participants. Trial outcome assessment
Fingerprint check. Screened for TB and HIV by Research Assistants in Study Research Office

Figure 1. PROSPECT Trial Design. WHO: World Health Organization, TB: tuberculosis, HIV: Human immunodeficiency virus, +ve: positive, -ve: negative, ART: antiretroviral therapy, Rx: treatment, +/- : with, or without.

Participants who test HIV-negative will be referred to the clinic waiting area with a copy of their HIV test result, and will be assessed by the routine facility health workers as in Group 1.

Participants whose oral fluid test is reactive will have confirmatory HIV testing performed by trained research assistants using rapid diagnostic kits and following a testing algorithm as recommended by WHO and Malawi National Guidelines.

Participants with confirmed HIV infection will be supported to register for HIV care at the HIV care clinic located within the study health centre and will be provided with a written appointment date to attend for initial antiretroviral therapy assessment

and initiation appointment. As part of HIV treatment assessments, Malawi National guidelines recommend screening for active tuberculosis. In this group, participants' TB screening will be directed by facility HIV clinic health workers - not by the study team - with sputum smear microscopy and GeneXpert MTB/Rif available to clinicians through routine clinic services, and without further study intervention.

Group 3: Optimised TB and HIV screening and linkage to care
Participants allocated to Group 3 will be directed to a research study room located in a separate area of the clinic. Prior to intervention delivery, a digital fingerprint check will be repeated to validate participants' identity and minimise any potential

for contamination between groups. Participants will be offered supervised HIV self-testing, with confirmatory testing and post-test counselling as in Group 2 above.

All participants allocated to Group 3 will additionally be offered a digital chest x-ray (MinXray Inc., USA) that will be pre-processed and quality-checked by a study radiographer. Digital chest x-ray images will then be immediately evaluated by the CAD4TB computer-aided TB triage algorithm (Delpht Imaging Systems, Netherlands). Application of the CAD4TB algorithm will provide a TB score ranging from 0 to 100. Based on a pilot evaluation of chest x-rays taken from adults attending the study clinic with cough and analysed by Delpht Imaging services using CAD4TB version 4.12.2, we set a threshold CAD4TB score of 45. Participants with a CAD4TB score at or above this threshold will be classified as having "high probability of tuberculosis", whilst those below this threshold will be classified as having "low probability of TB".

Participants whose chest x-rays are classified as having "high probability of TB" will be invited to submit a single sputum sample (induced with saline nebulization if necessary) for testing by GeneXpert MTB/Rif (Cephaid, USA) in the study clinic. Participants whose GeneXpert MTB/Rif results demonstrate the presence of *M. tuberculosis* will be supported to register for tuberculosis treatment on the same day at the TB clinic within the study health facility.

Participants whose chest x-ray are classified as being "low probability of TB" and those whose GeneXpert MTB/Rif tests are negative will be referred to the clinic waiting room to be seen by facility health workers, with a written report of the results of the investigations they have completed.

Definitions

Microbiologically-confirmed tuberculosis will be defined by: A participant with: a documented positive GeneXpert MTB/Rif result for *Mycobacterium tuberculosis* on at least one sample of sputum taken for study or routine clinical purposes; or documented growth of *Mycobacterium tuberculosis* and positive speciation using MPT 64 antigen tests on at least one culture of sputum taken for study or routine clinical purposes; or documented identification of acid fast bacilli on at least one sputum sample taken for study or routine clinical purposes and examined by sputum smear microscopy.

Clinically-diagnosed tuberculosis will be defined by: A participant who does not fulfil the criteria for microbiologically-confirmed TB but has documented evidence of having been diagnosed with active TB by a clinician or other medical practitioner who has decided to give the patient a full course of TB treatment. This definition includes cases diagnosed on the basis of X-ray abnormalities or suggestive histology and extrapulmonary cases without laboratory confirmation.

Clinically-diagnosed cases that are subsequently found to be bacteriologically-positive (before or after starting treatment) will be reclassified as bacteriologically-confirmed.

Pulmonary tuberculosis (PTB) will be defined by: A participant with bacteriologically-confirmed or clinically diagnosed case of TB involving the lung parenchyma or the tracheo-bronchial tree. Miliary TB will be classified as PTB because there are lesions in the lungs. Tuberculous intra-thoracic lymphadenopathy (mediastinal and/or hilar) or tuberculous pleural effusion, without radiographic abnormalities in the lungs, constitutes a case of extrapulmonary TB. A patient with both pulmonary and extrapulmonary TB will be classified as a case of PTB.

Extrapulmonary tuberculosis (EPTB) will be defined by: A participant with bacteriologically-confirmed or clinically-diagnosed case of TB involving organs other than the lungs, e.g. pleura, lymph nodes, abdomen, genitourinary tract, skin, joints and bones, meninges.

Initiation of tuberculosis treatment will be defined by: A participant in whom there is documented evidence of commencement of anti-tuberculosis treatment, either by: inspection of the participant-carried national tuberculosis treatment card; or inspection of the facility tuberculosis treatment register; or inspection of TB treatment medication bottles or pill boxes.

Initiation of antiretroviral therapy will be defined by: A participant in whom there is documented evidence of commencement of combination antiretroviral therapy treatment, either by: inspection of the participant-carried national HIV programme treatment card; or inspection of the facility antiretroviral therapy treatment register; or inspection of antiretroviral therapy medication bottles or pill boxes.

Successful tuberculosis treatment outcome will be defined by: A participant in whom tuberculosis treatment is initiated for bacteriologically-confirmed pulmonary tuberculosis, and who has documented evidence in their national tuberculosis treatment card of being cured of TB, being either sputum smear- or culture-negative in their last month of treatment and on at least one previous occasion; or a participant with documented evidence of having completed TB treatment without evidence of failure (that is sputum smear- or culture-positive at month 5 or later during treatment) but with no record to show that sputum smear or culture results in the last month of treatment and on at least one previous occasion were negative, either because tests were not done or because results are unavailable.

Adverse events

As this is a pragmatic randomised trial and no new investigational products are being evaluated, we anticipate only a

small number of adverse events. Nevertheless, we will ensure that case definitions, standardised operating procedures and a reporting protocol will be in place to record all adverse events.

The following adverse events will be systematically recorded and reported:

- Misclassification or misinterpretation of results leading to a participant starting TB therapy in error

- Misclassification or misinterpretation of results leading to a participant starting HIV treatment in error

- Breach of confidentiality following TB or HIV diagnosis

- Needlestick injuries

Trial outcomes

The primary trial outcome will be time in days – from Day 0 up to but not including Day 56 – to tuberculosis treatment initiation, evaluated at Day 56 following randomization.

Analysis of the primary outcome will be done on an intention to treat basis, with all participants allocated to trial groups included and analysed in the group to which they were randomized (regardless of which intervention was received). We will make three pair-wise comparisons *(Group 2 vs. Group 1; Group 3 vs. Group 2; and Group 3 vs. Group 1)*.

This primary endpoint has been chosen because reducing time to initiation of treatment could have important individual and public health benefits. Assessment over eight weeks has been selected because: (i) TB culture is typically completed within 8 weeks, (ii) mortality is highest during this period[27,28], and (iii) previous trials and our previous research show that TB treatment initiations plateau by 8 weeks[27].

The secondary trial outcomes will be:

- The proportion of randomised participants initiated onto tuberculosis treatment on the same day as randomisation, with the numerator being participants who were initiated on tuberculosis treatment on Day 0, and the denominator being all randomised participants.

- The proportion of randomised participants with undiagnosed/untreated microbiologically-confirmed pulmonary TB at Day 56, with the numerator being participants with microbiologically-confirmed tuberculosis (either sputum culture, or sputum Xpert, or sputum smear microscopy positive on a sample taken on Day 56) and who are confirmed not to be taking tuberculosis treatment on Day 56 (including participants who have previously initiated tuberculosis treatment, but have defaulted or stopped treatment – regardless of reason – for at least one week). The denominator will be all randomised participants.

- The proportion of randomised participants with undiagnosed/untreated HIV at Day 56, with the numerator being participants with positive confirmatory HIV test results at Day 56 and who are not taking antiretroviral

therapy (regardless of previous HIV test results during or before the study period), and the denominator being all randomised participants.

- Time in days - from Day 0 up to but not including Day 56 - to initiation of antiretroviral therapy among participants with positive confirmatory HIV test results at Day 56 and who were not taking antiretroviral therapy at Day 0.

- The proportion of randomised participants reported to have died by Day 56, with the numerator being participants confirmed to have died through home tracing visits or TB treatment records, and the denominator being all randomised participants

- The proportion of TB cases with a successful TB treatment outcome. The numerator will be participants who were initiated onto tuberculosis treatment (either microbiologically-confirmed or clinically-diagnosed tuberculosis) up to, but not including Day 56, and who have a successful TB treatment outcome (either cured or completed treatment) at 6-months after starting treatment. The denominator will be all participants confirmed to have initiated tuberculosis treatment between Day 0 and up to, but not including Day 56.

- Mean difference in EuroQoL EQ5D utility score at Day 56, adjusting for participants' EQ5D utility score measured at Day 0.

- Mean difference in EuroQoL EQ5D visual analogue scale score, adjusting for participants' EQ5D visual analogue scale score measured at Day 0.

- Incremental cost-effectiveness per quality-adjusted life year gained

Planned subgroup analyses

In pre-planned exploratory analysis, we will stratify analysis of the primary outcome by: sex (male vs. female); and microbiological status (bacteriologically-confirmed TB vs. clinically-diagnosed TB).

Additionally, we will undertake a Bayesian analysis of the primary trial outcome[29]. Prior distributions for the proportion initiating TB treatment under each intervention strategy will be elicited from key stakeholder groups. We anticipate that key stakeholders will include: community members; clinic health workers; researchers; TB/HIV experts; and policymakers (Malawi, regional, and international). Before eliciting stakeholders' prior beliefs for trial interventions, we will provide a series of "warm-up" vignettes based around familiar events such as the probability of a football team winning, or the probability of it raining tomorrow. Each stakeholder will then be asked to make ten guesses for the percentage of participants who will initiate TB treatment under each intervention strategy.

Outcome evaluation

Following completion of trial interventions, all participants in each of the three groups will be given a written appointment

card to attend a follow-up assessment at the study research clinic room 56 days after randomization (or as close as possible after this date). They will also be issued with a voucher that they can use to reimburse the cost of transport to the clinic for this assessment. Participants who don't attend their day 56 appointment will be traced to home.

To evaluate the primary trial outcome (time to tuberculosis treatment initiation), Research Assistants will undertake a detailed questionnaire to record the date of TB treatment initiation, and verify by inspecting participant-carried TB treatment cards, medication and clinic TB treatment registers.

To evaluate the prevalence of undiagnosed tuberculosis, Research Assistants will collect sputa from all participants. Samples will be transported to the TB Research Laboratory at the College of Medicine of Malawi, where they will be cultured for tuberculosis using the MGIT system, undergo smear examination using fluorescence microscopy, and tested using the GeneXpert MTB/Rif assay. Positive TB results will be reported to participants within three days of receipt (including by home tracing), and participants will be supported to register for TB treatment at the health facility. Participants with Rifampicin resistance detected on GeneXpert will be traced and supported to access the TB clinic at the Queen Elizabeth Central Hospital for further clinical assessment and evaluation for treatment.

To evaluate the prevalence of undiagnosed HIV at Day 56, Research Assistants will offer all participants HIV testing, unless they are confirmed to be taking antiretroviral therapy. All participants requiring additional care will be supported to access either the HIV clinic, TB clinic, or outpatient clinic at the study clinic as required. Additionally, we will support referral and access to Queen Elizabeth Central Hospital should further specialist care be required.

To evaluate health-related quality of life (HRQoL), we will use the EQ-5D-3L. The EQ-5D is a generic HRQoL measure, and was translated into Chichewa following international and EuroQoL guidelines. The EQ-5D-3L tool will be administered to all trial participants at baseline, and on Day 56.

Statistical methods

The primary trial outcome, time to TB treatment initiation, will be compared between pairs of groups among all randomised participants. To evaluate the relative effects of the HIV and TB screening/linkage interventions, we will make three pairwise comparison *(Group 2 vs. Group 1; Group 3 vs Group 2; and Group 3 vs Group 1)*. Our pilot data show that 17% of adults with TB symptoms will initiate TB treatment under routine screening conditions within 8-weeks.

Using formula for the proportional hazards model developed by Schoenfield[30], and inflating by 5% for loss to follow-up, a total sample size of 1455 participants (485 per group) gives at least 80% power to detect at least a cumulative hazard ratio for TB treatment initiation of 1.5 comparing Group 2 to Group 1, and a hazard ratio of 1.41 comparing Group 3 to Group 2, at 5%

significance level. Additionally, under these assumptions, 485 participants per group would give 80% power to detect a hazard ratio of at least 1.50 comparing Group 3 to Group 1.

All statistical analysis will be conducted in accordance with a pre-published statistical analysis plan (Supplementary File 2). Trial reporting will follow CONSORT Guidelines. We will report baseline characteristics of randomised participants, stratified by allocated group.

Analysis of the primary and secondary outcomes will be done on an intention to treat basis, with all participants allocated to trial groups included. We will index the day of recruitment to be Day 0 and outcome assessment will take place on, or as close to possible after, Day 56. Initiation of tuberculosis treatment will be defined by a participant in whom there is documented evidence of commencement of anti-tuberculosis treatment between Day 0 and up to, but not including Day 56. Time to TB treatment outcome analysis will be right-censored on day 56 if TB treatment is not initiated. We will estimate per-group median times to TB treatment initiation, and plot cumulative hazard function graphs.

To investigate the relative effectiveness of interventions on the cumulative hazard of TB treatment initiation, we will conduct log rank tests and construct Cox proportional hazard regression models to estimate hazard ratios and 95% confidence intervals for each pairwise comparison *(e.g. Group 2 vs. Group 1, Group 3 vs. Group 2, and Group 3 vs. Group 1)*. Log-log plots will be examined and Schoenfeld residuals used to test the proportional hazards assumption.

To analyse binary secondary outcomes (proportion with same-day tuberculosis treatment initiation, proportion with undiagnosed/untreated pulmonary tuberculosis, proportion with undiagnosed/untreated HIV, proportion reported to have died by Day 56, proportion with successful TB treatment outcome), we will construct log-binomial regression models to estimate relative risk ratios and 95% confidence intervals, comparing between pairs of groups. We will additionally compare between pairs of groups the time to antiretroviral therapy initiation among participants with previously untreated HIV using Cox regression models.

To evaluate the effect of interventions on health-related quality of life, we will use ANCOVA analysis to compare the mean EQ5D utility scores and visual analogue scale scores measured at Day 56 between pairs of groups, adjusting for participants' values measured at Day 0.

For the preplanned subgroup analysis of the primary trial outcome we will construct Cox proportion hazard regression models including a term for either sex (male or female) or microbiological TB status (either microbiologically confirmed or clinically-diagnosed) to estimate hazard ratios and 95% confidence intervals. We will use the likelihood ratio test to look for interactions between sex/microbiological-confirmed TB and trial group.

Bayesian analysis of primary trial outcome

Using within and between participant elicited probability distributions, we will construct stakeholder group-specific pooled prior probability distributions (known as a "community of priors"). Each prior will be converted to a log-hazard ratio scale and fitted to a normal distribution, allowing comparison between stakeholder groups of the similarity in support of opinions of effectiveness and of uncertainty.

Using Bayes' theorem we will combine elicited stakeholder group-specific log hazard ratio prior distributions with log-likelihood hazard ratio distributions from each pairwise comparison being made in the PROSPECT Study to construct posterior probability distributions. All analysis will be done in R and posterior mean hazard ratios and 95% credible intervals will be estimated by taking draws from the posterior distributions using the No-U-Turn Sampler (NUTS) implemented with Stan.

Additional nested analysis

WHO has recommended that *"Computer aided diagnosis can be used for TB detection for research, ideally following a protocol that contributes to the required evidence base for guideline development"*[22]. The PROSPECT Study therefore offers opportunity to undertake a nested evaluation to contribute to the evidence base for the diagnostic accuracy (as well as effectiveness) of the CAD4TB platform.

Participants for this nested evaluation will be adults recruited to the main PROSPECT Study trial, and who complete a Day 56 outcome TB screening assessment. As part of this outcome assessment, all participants will undergo CAD4TB classification of digital chest x-ray, as well as sputum testing by GeneXpert, and TB liquid automated culture. All digital chest x-rays taken from participants at outcome assessment will be uploaded to the password protected and secure MinXray online picture archiving and communication (PACS) radiology cloud server. All participant identifiers will be removed from x-rays prior to upload.

A panel of seven radiologists will each - independently and blinded to participant characteristics, HIV status, and results of microbiological investigations - classify chest x-rays using a standardised form for classification of chest radiology findings. Radiologists will review chest radiographs, and using an online data entry form, indicate the presence of any:

- Infiltrate or consolidation

- Cavitary lesion

- Nodule or mass with poorly defined margins

- Hilar/mediastinal adenopathy

- Pleural effusion

- Milliary findings

- Discrete linear opacity

- Discrete nodule(s) without calcification

- Other findings

Radiologists will additionally classify chest radiographs as either suggestive of active pulmonary tuberculosis, or not. We will use the kappa statistic (two outcomes, multiple readers) with 95% confidence intervals to assess inter-reader agreement among the radiologists.

For the diagnostic accuracy evaluation, the index test will be CAD4TB score (continuous variable ranging from 0 to 100, and in a secondary analysis, dichotomised for greater than or equal to the CAD4TB threshold score of 45).

Reporting will follow the STARD Guidelines. For each index test definition (continuous distribution, and dichotomised to high vs. low probability of tuberculosis), we will compare diagnostic accuracy against two pre-defined reference standards:

1) Consensus radiologist classification with at least 5/7 independent readers agreeing that the radiograph was "suspicious of tuberculosis" (with sensitivity analysis limited to cases only where all 7/7 readers agreed), and

2) Bacteriologically-confirmed pulmonary tuberculosis, defined as either a documented positive GeneXpert result for *Mycobacterium tuberculosis* on at least one sample of sputum taken for study purposes at Day 56 assessment; or documented growth of *Mycobacterium tuberculosis* and positive speciation using MPT 64 antigen tests on at least one culture of sputum taken for study purposes at Day 56 assessment; or documented identification of acid fast bacilli on at least one sputum sample taken for study purposes and examined by sputum smear microscopy at Day 56 assessment.

For each comparison, sensitivity, specificity, positive predictive value, negative predictive value, and diagnostic odds ratios will be reported. Additionally, by constructing logistic regression models, we will investigate the effect of reader characteristics (practicing in Africa or elsewhere, years of practice) on diagnostic accuracy with bacteriological-confirmation as the reference standard.

Validation of urinary lipoarabinomannan testing (LAM)

Urine LAM testing is a relatively new tuberculosis diagnostic, that has high accuracy among adults with advanced HIV infection, and has been shown to reduce mortality in hospitalised HIV-positive adults[9]. The test is based on a lateral flow assay, and has been constructed to be used as a point of care test, with results read at the bedside.

However, sensitivity is known to be suboptimal among ambulant TB suspects. A newer version of the urine LAM test (FIND/Fujifilm) has been reported to have high sensitivity for TB, even among ambulant adults, and HIV-negative individuals. Therefore, the PROSPECT Study offers opportunity to undertake a nested evaluation of the performance of this test. We will collect a 5ml sample of urine from all participants at baseline, and transport the sample to the TB laboratory at the

College of Medicine of Malawi for urine LAM testing. We will compare the diagnostic yield of urine LAM testing with that of sputum culture, smear and Xpert from day 56 participant samples.

Data handling and management

Data will be collected by research assistants using the mobile CommCare data collection platform running on fingerprint secured tablets. Data will be transmitted to the secure study server over encrypted cellular networks. The MLW Data Department has considerable experience in building robust electronic data collection surveys and in secure data management, backup and processing. A full audit trial of database changes will be maintained.

Building upon our extensive experience of conducting previous trials using electronic data collection systems in Blantyre, the trial statistician and Chief Investigator will write scripts within the statistical programme R that will interface with the trial database and, on a regular automated basis, use logical rules to identify records with missing or implausible values that will be hand-checked against source records to ensure completeness and validity of the final dataset.

We are strongly committed to ensuring that the trial datasets are made openly available, and that all code used in the analysis are published to allow fully reproducible research. The data collected by this research will be of importance to other researchers and the public, and could for example be used by other researchers conducting meta-analysis, or by policymakers modelling the potential return on investment of implementing interventions within their settings. Therefore, we will establish a public online GitHub repository, where the final anonymised individual- level trial dataset and code to allow reproduction of all analysis will be published. The availability of these resources will be publicised within academic manuscripts, through the MLW and LSTM websites.

Economic analysis

Two economic evaluations will be undertaken: firstly, a within-trial evaluation; and secondly, a decision-analytic based cost effectiveness model. Both will be used to estimate the expected incremental cost per quality-adjusted life year (QALY) gained for the two optimized TB/HIV interventions in comparison to standard of care. For both analyses, the perspective will be that of the Malawi Ministry of Health, and will only include the direct medical costs.

The within trial evaluation will adopt a time horizon matching the length of follow-up in the trial. The model-based evaluation will adopt a lifetime horizon so as to incorporate the long-term costs and health consequences of delayed TB/HIV diagnosis and treatment initiation.

For the within-trial evaluation, total costs and health benefits (QALYs) will be calculated over 56 days for each participant in each trial arm. Healthcare resource utilisation (e.g. clinic visits; investigations; medications) will be recorded over the 56 days from randomisation. Unit costs for these healthcare resources

will be derived from primary costing studies, previous costing studies in Malawi or from targeted literature searches and inflated to the year of analysis. Unit costs for medications will be taken from the Management Sciences for Health International Drug Price Indicator guide. Responses to the EQ-5D-3L will be converted to health state utility values using the Zimbabwean tariff set and combined with the time spent within each health state to generate QALYs.

As the distributions of costs and QALYs are commonly skewed, and often bimodal or truncated, a range of estimators will be explored, and model diagnostics will be undertaken to determine optimal choice. Mean costs and outcomes for each intervention will be estimated, together with the mean incremental cost-effectiveness ratio. Measures of uncertainty (standard errors and confidence intervals) will also be reported for the mean estimates. The ICER will be calculated by comparing the least costly trial arm to the next least costly arm and calculated as below:

$$\text{ICER} = (\text{Cost}_2 - \text{Cost}_1)/(\text{QALY}_2 - \text{QALY}_1)$$

Alternative trial arms that are more costly and less effective will be interpreted as dominated and would not represent an efficient use of resources. For alternative trial arms that are more costly and more effective, the interpretation of cost-effectiveness depends on the policy makers willingness to pay threshold (WTP) for a gain in QALY. Malawi and the majority of other African countries do not have an explicit WTP thresholds for interpreting cost-effectiveness. Hence cost-effectiveness acceptability curves (CEACs) will be constructed to identify the optimal intervention at different WTP thresholds.

The model-based evaluation will aim to extrapolate trial findings to allow estimation of cost-effectiveness over a lifetime time horizon. The model will likely consist of mutually exclusive Markov health states. These health states will be defined by a combination of untreated and treated states for both HIV and TB. The model will be parametrised by findings observed in the trial, and data extracted from the published literature.

Translating research into policy

Results of this research will be important in guiding national, regional and international health policy. WHO, policymakers and parliamentarians are currently grappling with how to improve access to TB and HIV diagnosis and treatment, including the role of chest x-ray. A key objective of this study is therefore to translate research findings into normative guidance in Malawi, in sub-Saharan Africa, and through WHO.

We recognise that early engagement with policymakers is essential to translate research into action. Therefore, we undertaken preliminary scoping activities to identify key stakeholders that we will work with, including Malawi Ministry of Health TB/HIV Technical Working Groups, the Malawi Network for Evidence-Informed Decision Making (EvIDeNt) which includes regional linkage through the African Institute for Development Policy (AFIDEP), and WHO TB-STAG.

Ethical considerations

This research has been approved by the College of Medicine of Malawi Research Ethics Committee (COMREC – number: P.11/17/2311), and the Research Ethics Committee of the Liverpool School of Tropical Medicine (number: 17-050). Trial progress will be reviewed by a data and safety monitoring board.

All prospective participants will be asked to provide written informed consent to take part in the trial. Individuals who are illiterate will be asked to provide a witnessed thumbprint to confirm their informed consent to participate. Witnesses will be an independent individual not involved with the study.

Timelines

Piloting and preparatory activities will commence in April 2018, and trial recruitment in November 2018. We anticipate recruiting participants over an 8-month period, with final outcome assessment of treatment outcomes and mortality conducted after 6-months. Thus, the study will be completed in August 2020.

Trial registration

This trial was registered with ClinicalTrials.gov on the 8th May 2018 (NCT03519425).

Discussion

The PROSPECT Study will use a pragmatic trial design[31] to evaluate optimised TB/HIV screening and treatment linkage interventions under "real-life" condition in primary care in Malawi. The three-arm design allows us to efficiently test two important hypotheses.

Firstly, by comparing Group 2 to Group 1, we will investigate whether HIV care should be prioritised for adults with symptoms of TB. In previous studies, we and others have found operationalising universal HIV testing for adults attending primary care challenging due to limited counsellor capacity[8], meaning that only individuals in whom clinicians had a high suspicion of HIV were referred to the counsellor for HIV testing[32]. Implementing a semi-supervised HIV self-testing intervention could then free-up counsellor capacity, increasing testing coverage. Moreover, self-testing is popular with patients, as it allows them to take control of the manner in which they learn their HIV status[25,33]. In addition to ensuring that individuals with TB symptoms are not "caught" between the HIV and TB care clinics[16], facilitated linkage to HIV care may promote earlier, more intensive screening for TB than would have otherwise occurred in the general outpatient clinic. Initiation of antiretroviral therapy may also unmask TB in individuals with advanced immunosuppression[34], prompting an earlier clinical decision to initiate tuberculosis treatment.

By comparing Group 3 to Group 2 and Group 1, we will provide strong evidence on the effectiveness of a novel triage approach to TB screening. As no single TB diagnostic currently has optimal characteristics in terms of test accuracy, reliability,

implementation and scalability at primary care level, and cost per case detected, a triage approach that comprises an efficient, high sensitivity initial test, followed by a high specificity confirmatory test is required for patients with symptoms of TB. WHO recently undertook a modelling exercise to compare potential triage testing approaches, and found that, among adults with TB symptoms, an algorithm comprised of an initial chest x-ray followed by a confirmatory GeneXpert MTB/Rif test would likely fulfil requirements of having a high overall sensitivity and specificity, low number needed to screen to detect a case, and low cost per case detected[22]. Moreover, this approach is attractive as, with the increasing availability of affordable digital x-ray units, the entire triage algorithm can be completed on the same day and within the same clinic. As an initial screening tool, chest x-ray can be performed quickly for large numbers of cases, screening out those with a low probability of disease. However, until now, widespread implementation has been limited by cost and by the availability of trained radiologists to classify x-rays. Therefore, in this study, we will evaluate the effectiveness of the CAD4TB computer aided diagnosis software system as the chest x-ray reader, which has demonstrated high accuracy (comparable to radiologists and clinicians) in diagnostic accuracy studies in Europe and Africa[23]. In a nested evaluation, we will examine the diagnostic accuracy of the CAD4TB system compared to a panel of radiologists and sputum culture for *Mycobacterium tuberculosis.*

There are some limitations to this study. Although we will take extensive steps to minimise contamination between groups, including by participant fingerprint validation prior to intervention delivery, undertaking the trial in a single primary health may influence routine care decisions made by clinicians. There is a possibility of a Hawthorne effect, which might reduce the size of the differences between the standard of care arm and the interventions arms, and thus power might be reduced. Should evidence for effectiveness of these interventions be found, further supportive evidence would be provided by a future trial that randomly allocated clinics to intervention groups, and by surveillance of key process indicators under routine implementation conditions. The primary trial outcome will compare the time to TB treatment initiation between groups, an important indicator of individual and public health effectiveness. However, future larger studies may wish to investigate effectiveness against mortality, here investigated as a secondary outcome. Finally, as a sputum sample for TB culture will not be taken from participants in all groups at baseline, we will not be able to estimate the effectiveness of interventions on participants with true microbiologically-confirmed disease. This was a deliberate design of the study; TB culture is not widely-available in Africa as standard of care, meaning that empirical TB treatment is common[35]. Should we have offered participants TB culture of sputum at baseline, we would not be able to obtain a true estimate of the effectiveness of interventions under "real-life" conditions. To ensure all participants are provided with high-quality screening and care, at Day 56 they will all be offered TB and HIV screening and supported to access care.

An important component of the PROSPECT Study is the evaluation of effects on patient-important outcomes – as measured by change in health-related quality of life – and cost-effectiveness. In Blantyre, we have well-established systems for patient and health resource costing[36], and will use both a within-trial cost-effectiveness evaluation, as well as a decision-analytic based cost effectiveness model. Additionally, we will work through the Policy Unit at the Malawi-Liverpool-Wellcome Trust Clinical Programme to engage early with national TB and HIV programmes in Malawi, and with regional and supranational policy fora. The exploratory Bayesian trial analysis will incorporate the prior beliefs of various groups of stakeholders to ensure that meaningful evidence can be provided to key groups.

Dissemination of findings

Trial results and findings from the cost-effectiveness analysis will be shared with Blantyre District Health Office, the Malawi National Tuberculosis Programme and with the Malawi National HIV Programme. We will report findings at national, regional and international conferences, and will submit a manuscript reporting trial findings to a peer-reviewed journal specialising in public health, HIV and tuberculosis.

To facilitate reproducibility of analysis, an anonymised minimal final dataset and all code required to reproduce analysis will be published in the trial GitHub repository.

Supplementary material
Supplementary File 1 – Baseline questionnaire.

Click here to access the data.

Supplementary File 2 – Statistical analysis plan.

Click here to access the data.

Current trial status
The trial is currently in preparatory and piloting phase.

In summary, the PROSPECT Study will provide urgently-needed evidence under "real-life" conditions to inform clinicians and policy makers on how best to improve TB/HIV diagnosis and treatment initiation in Africa.

Data availability
All data underlying the results are available as part of the article and no additional source data are required.

References

1. World Health Organization: **Global tuberculosis report 2017.** Geneva, 2017.
   Reference Source

2. Murray CJ, Ortblad KF, Guinovart C, *et al.*: **Global, regional, and national incidence and mortality for HIV, tuberculosis, and malaria during 1990-2013: a systematic analysis for the Global Burden of Disease Study 2013.** *Lancet.* 2014; **384**(9947): 1005–70.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

3. Corbett EL, Watt CJ, Walker N, *et al.*: **The growing burden of tuberculosis: global trends and interactions with the HIV epidemic.** *Arch Intern Med.* 2003; **163**(9): 1009–21.
   PubMed Abstract | **Publisher Full Text**

4. World Health Organization: **END TB Strategy.** Geneva, 2016.
   Reference Source

5. Khan MS, Fletcher H, London School of Hygiene and Tropical Medicine TB Centre Steering Committee, *et al.*: **Investments in tuberculosis research - what are the gaps?** *BMC Med.* 2016; **14**(1): 123.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

6. Harries AD, Lawn SD, Getahun H, *et al.*: **HIV and tuberculosis--science and implementation to turn the tide and reduce deaths.** *J Int AIDS Soc.* 2012; **15**(2): 17396.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

7. Gupta RK, Lucas SB, Fielding KL, *et al.*: **Prevalence of tuberculosis in post-mortem studies of HIV-infected adults and children in resource-limited settings: a systematic review and meta-analysis.** *AIDS.* 2015; **29**(15): 1987–2002.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

8. MacPherson P, Lalloo DG, Choko AT, *et al.*: **Suboptimal patterns of provider initiated HIV testing and counselling, antiretroviral therapy eligibility assessment and referral in primary health clinic attendees in Blantyre, Malawi.** *Trop Med Int Health.* 2012; **17**(4): 507–17.
   PubMed Abstract | **Publisher Full Text** | **Free Full Text**

9. Peter JG, Zijenah LS, Chanda D, *et al.*: **Effect on mortality of point-of-care, urine-based lipoarabinomannan testing to guide tuberculosis treatment initiation in HIV-positive hospital inpatients: a pragmatic, parallel-group, multicountry, open-label, randomised controlled trial.** *Lancet.* 2016; **387**(10024):

1187–97.
PubMed Abstract | **Publisher Full Text**

10. MacPherson P, Corbett EL, Makombe SD, *et al.*: **Determinants and consequences of failure of linkage to antiretroviral therapy at primary care level in Blantyre, Malawi: a prospective cohort study.** *PLoS One.* 2012; **7**(9): e44794.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

11. Macpherson P, Dimairo M, Bandason T, *et al.*: **Risk factors for mortality in smear-negative tuberculosis suspects: a cohort study in Harare, Zimbabwe.** *Int J Tuberc Lung Dis.* 2011; **15**(10): 1390–6.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

12. MacPherson P, Houben RM, Glynn JR, *et al.*: **Pre-treatment loss to follow-up in tuberculosis patients in low- and lower-middle-income countries and high-burden countries: a systematic review and meta-analysis.** *Bull World Health Organ.* 2014; **92**(2): 126–38.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

13. World Health Organization: **Systematic screening for active tuberculosis: principles and recommendations.** Geneva, 2013.
PubMed Abstract

14. Cazabon D, Alsdurf H, Satyanarayana S, *et al.*: **Quality of tuberculosis care in high burden countries: the urgent need to address gaps in the care cascade.** *Int J Infect Dis.* 2017; **56**: 111–116.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

15. Chihota VN, Ginindza S, McCarthy K, *et al.*: **Missed Opportunities for TB Investigation in Primary Care Clinics in South Africa: Experience from the XTEND Trial.** *PLoS One.* 2015; **10**(9): e0138149.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

16. MacPherson P, Choko AT, Webb EL, *et al.*: **Development and validation of a global positioning system-based "map book" system for categorizing cluster residency status of community members living in high-density urban slums in Blantyre, Malawi.** *Am J Epidemiol.* 2013; **177**(10): 1143–7.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

17. Davis JL, Cattamanchi A, Cuevas LE, *et al.*: **Diagnostic accuracy of same-day microscopy versus standard microscopy for pulmonary tuberculosis: a systematic review and meta-analysis.** *Lancet Infect Dis.* 2013; **13**(2): 147–54.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

18. Steingart KR, Henry M, Ng V, *et al.*: **Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review.** *Lancet Infect Dis.* 2006; **6**(9): 570–81.
PubMed Abstract | **Publisher Full Text**

19. Cruciani M, Scarparo C, Malena M, *et al.*: **Meta-analysis of BACTEC MGIT 960 and BACTEC 460 TB, with or without solid media, for detection of mycobacteria.** *J Clin Microbiol.* 2004; **42**(5): 2321–5.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

20. Steingart KR, Schiller I, Horne DJ, *et al.*: **Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults.** *Cochrane Database Syst Rev.* 2014; (1): Cd009593.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

21. Hoog AH, Meme HK, van Deutekom H, *et al.*: **High sensitivity of chest radiograph reading by clinical officers in a tuberculosis prevalence survey.** *Int J Tuberc Lung Dis.* 2011; **15**(10): 1308–14.
PubMed Abstract | **Publisher Full Text**

22. World Health Organization: **Chest Radiology in Tuberculosis Detection: Summary of recommendation and guidance on programmatic approaches.** Geneva, 2016.
Reference Source

23. Pande T, Cohen C, Pai M, *et al.*: **Computer-aided detection of pulmonary tuberculosis on digital chest radiographs: a systematic review.** *Int J Tuberc Lung Dis.* 2016; **20**(9): 1226–30.
PubMed Abstract | **Publisher Full Text**

24. National Tuberculosis Programme Malawi: **National Tuberculosis Prevalence Survey.** 2015.

25. Choko AT, MacPherson P, Webb EL, *et al.*: **Uptake, Accuracy, Safety, and Linkage into Care over Two Years of Promoting Annual Self-Testing for HIV in Blantyre, Malawi: A Community-Based Prospective Study.** *PLoS Med.* 2015; **12**(9): e1001873.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

26. Kumwenda M, Desmond N, Hart G, *et al.*: **Treatment-Seeking for Tuberculosis-Suggestive Symptoms: A Reflection on the Role of Human Agency in the Context of Universal Health Coverage in Malawi.** *PLoS One.* 2016; **11**(4): e0154103.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

27. Nliwasa M, MacPherson P, Chisala P, *et al.*: **The Sensitivity and Specificity of Loop-Mediated Isothermal Amplification (LAMP) Assay for Tuberculosis Diagnosis in Adults with Chronic Cough in Malawi.** *PLoS One.* 2016; **11**(5): e0155101.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

28. Auld AF, Fielding KL, Gupta-Wright A, *et al.*: **Xpert MTB/RIF - why the lack of morbidity and mortality impact in intervention trials?** *Trans R Soc Trop Med Hyg.* 2016; **110**(8): 432–44.
PubMed Abstract | **Publisher Full Text**

29. Speigelhalter D, Abrams K, Myles J: **Bayesian approaches to Clinical trials and health-care evaluation.** Chichester, England: John Wiley and Sons ltd.; 2004.
Publisher Full Text

30. Lachin JM: **Sample size and power for a logrank test and Cox proportional hazards model with multiple groups and strata, or a quantitative covariate with multiple strata.** *Stat Med.* 2013; **32**(25): 4413–25.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

31. Sox HC, Lewis RJ: **Pragmatic Trials: Practical Answers to "Real World" Questions.** *JAMA.* 2016; **316**(11): 1205–6.
PubMed Abstract | **Publisher Full Text**

32. MacPherson P, MacPherson EE, Mwale D, *et al.*: **Barriers and facilitators to linkage to ART in primary care: a qualitative study of patients and providers in Blantyre, Malawi.** *J Int AIDS Soc.* 2012; **15**(2): 18020.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

33. Choko AT, Fielding K, Stallard N, *et al.*: **Investigating interventions to increase uptake of HIV testing and linkage into care or prevention for male partners of pregnant women in antenatal clinics in Blantyre, Malawi: study protocol for a cluster randomised trial.** *Trials.* 2017; **18**(1): 349.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

34. Lai RP, Meintjes G, Wilkinson RJ: **HIV-1 tuberculosis-associated immune reconstitution inflammatory syndrome.** *Semin Immunopathol.* 2016; **38**(2): 185–98.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

35. Theron G, Peter J, Dowdy D, *et al.*: **Do high rates of empirical treatment undermine the potential effect of new diagnostic tests for tuberculosis in high-burden settings?** *Lancet Infect Dis.* 2014; **14**(6): 527–32.
PubMed Abstract | **Publisher Full Text**

36. Maheswaran H, Petrou S, MacPherson P, *et al.*: **Cost and quality of life analysis of HIV self-testing and facility-based HIV testing and counselling in Blantyre, Malawi.** *BMC Med.* 2016; **14**: 34.
PubMed Abstract | **Publisher Full Text** | **Free Full Text**

Open Peer Review

Current Peer Review Status: ✓ ✓

Version 3

Reviewer Report 26 November 2018

https://doi.org/10.21956/wellcomeopenres.16287.r34293

✓   **Tom A. Yates** (iD)

Section of Infectious Diseases and Immunity, Imperial College London, London, UK

Apologies for my delay in responding to these revisions. It is important that protocols are published before trials are undertaken and this document should now be indexed on databases. A few final thoughts below.

First, the authors are to be congratulated on undertaking a randomised evaluation of a carefully considered intervention with the potential to make a significant impact on avoidable morbidity and mortality. Robust evaluations of such programatic interventions (which usually means randomisation) are both challenging and incredibly important.

Thought will need to be exercised in interpreting the results of this trial, when it reports.

In the event of a null result, the possibility that the trial intervention has improved outcomes in the standard of care arm will need to be considered. The authors should think now about whether there are process data that might be collected to help understand a null result. Sophisticated analyses of null trial results can be instructive – for example, the work that was done to explain the results of the Thibela trial.

I agree that the proportion of participants at day 56 with undiagnosed microbiologically confirmed TB is a key outcome. I also agree that, in this context, under treatment is far riskier than over treatment. I still think that the trial would be stronger were baseline sputum samples collected from (at least) those initiating treatment at the initial study visit. This might be done, without introducing bias, at an exit interview. I appreciate that there are considerable resource implications but these data would assist in understanding the extent of overtreatment, an outcome of considerable importance for patients.

Anyway, enough from me. Good luck with this important study. I very much look forward to seeing the results.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 2

Reviewer Report 02 October 2018

https://doi.org/10.21956/wellcomeopenres.16035.r33594

✔ **Frank Cobelens**
[1] Department of Global Health, Academic Medical Center, Amsterdam, The Netherlands
[2] Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands

I thank the authors for their responses to the comments and revisions to the manuscript. I find all comments well addressed and approve the present version for indexing.

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Version 1

Reviewer Report 22 June 2018

https://doi.org/10.21956/wellcomeopenres.15894.r33288

? **Tom A. Yates** (iD)

Section of Infectious Diseases and Immunity, Imperial College London, London, UK

Thanks for asking me to review this protocol. I have read the protocol and the thoughtful peer review by Frank Cobelens. I have not had an opportunity to read the supplementary material. I comment as a clinician with training in epidemiology. The editors may wish to request a statistical

review and the views of a health economist.

I have divided my comments into those focused on the trial design and those focused on the description of the proposed study in the protocol.

Comments on the design of the trial
- ○ I share Frank's concerns about the potential for the trial interventions to alter clinical practice in the standard of care arm. The authors acknowledge this possibility and state that, should the intervention appear effective in this trial, they might proceed to a cluster RCT, randomising clinics. I would note that the Hawthorne Effect here would be more likely to bias the effect estimate towards the null, rather than lead to a spurious positive result.

- ○ I think Frank's comment about the primary outcome measure not capturing the number of people initiating TB treatment is an important one. I would add that, ideally, the primary outcome measure should consider whether initiation of TB treatment was appropriate or not. Empirical TB treatment may be more common in the standard of care arm, where clinicians may have less ready access to investigations. Early initiation of TB treatment in people who do not have TB is not a good thing. I would like to see more discussion about this possibility.

- ○ Not having cultures from all participants at baseline will make interpretation of the results difficult, particularly if there is a lot of empirical TB treatment (as I expect there may be). A good number of these patients may not have TB and unpicking that at day 56, by which time true cases of TB may have culture converted, may not be possible. Possible approaches to dealing with this issue...

  a) Take sputum for culture from everyone at baseline but don't feed the result back to clinicians until day 56. Is this morally less acceptable than deliberately not obtaining a sputum specimen? I would point out that delaying sputum collection until day 56 would lead to treatment initiation being further delayed in patients whose smear negative/Xpert negative TB had not already been diagnosed.

  b) Request that a baseline sputum sample for culture be taken from everyone in whom a decision to initiate TB treatment is made.

- ○ It would be good to see much more detail about what will happen during the day 56 study visit. If systematic recording of the adverse outcomes is to be achieved (especially 'Misclassification or misinterpretation of results leading to a participant starting TB therapy in error'), then presumably some determinantion will need to be made about which patients truly have TB and which do not. More detail about how this will be done should be included in the protocol, including how it will be done in patients that die before day 56.

- ○ It seems likely that clinicians making this assessment will need to speak to the patient (perhaps their relatives or clinicians who looked after them in hospital where patients have died) and to consult clinic notes from the first study visit. I struggle to see how this can be done whilst blinding the clinician to the group to which the patient was randomised, as is suggested will happen in the protocol.

○ I would suggest tempering the comments about this study being pragmatic. It strikes me that the trial is somewhere on the continuum between an efficacy trial and a truly pragmatic trial. Were this a truly pragmatic trial, the trial interventions would be implemented by existing clinic staff. Here, study staff implement the interventions. I would like the authors to comment on how they will unpick the impact of the trial interventions from the impact of the additional human resources available to patients randomised to Groups 2 and 3. It seems possible (even likely) that the trial interventions would be less effective if they were being implemented by busy clinic nurses?

○ I am not a health economist but wonder whether any differences in EQ5D scores might be diluted by including in this comparison lots of people who don't have TB (in whom the impact of the intervention may be minimal).

○ Frank's comment about non TB findings on chest x-ray is an important one. Might this trial be a good opportunity to quantify the burden of non TB disease that might be missed and whether that varies depending on whether clinicians have access to CAD?

○ With regards the second secondary outcome, I am unclear as to the rationale for excluding people that have stopped TB treatment for good reason (e.g. hepatotoxicity)?

○ With regards the nested evaluation of the CAD4TB tool, gold standard definitions of TB have been defined. Do gold standard definitions of 'not TB' need to also be defined, to allow estimates of specificity? I am not clear what is meant by 'reader characteristics' in the final paragraph of this section.

○ People with previous TB make up a significant proportion of those with TB disease in high burden settings (Florian Marx has done some nice work describing this phenomenon). These people may have residual x-ray changes and there are reports of false positive Xpert tests in individuals who have been recently treated. Will excluding people that have had TB within the preceding six months deal with the latter issue, or can DNA persist for longer than this in people with recent successfully treated TB?

○ I would like some justification for the Bayesian analysis to be included in the protocol. What is this achieving that couldn't be achieved using a frequentist approach?

○ I really like the open approach to sharing data and code that is proposed by the investigators.

Comments on the trial protocol
○ It seems like enrollment will precede randomisation but I would like to see a more explicit comment on allocation concealment.

○ I would like to see more comment about how patients who transfer their care out or are lost to follow up will be handled in the analysis.

○ The CAD threshold that will be used should be pre-specified in the protocol.

○ A note should be made on the approach that will be taken to x-rays in women of

childbearing age.

○ A note should be made about the approach that will be taken to patients that have RpoB mutations detected by Xpert.

○ It seems unlikely that the trial team will be able to see all participants on day 56. I suggest that the team allow themselves a window of time in which to undertake these study visits.

○ Please include an explicit comment about whether/how the urine LAM result will be fed back to clinicians.

○ There are a number of grammatical errors and the writing, in places, could be clearer. The section on trial outcomes, in particular, could be better written.

Conclusions
The proposed two step approach to evaluating people with suspected pulmonary TB in settings such as these seems to be a good idea. It is good to see this proposal to evaluate the approach in a randomised controlled trial. My main concerns about the protocol concern the choice of primary outcome, particularly how this might perform in a setting where clinicians often initiate empirical TB treatment. An attempt to ascertain 'true TB status' would be valuable in interpreting the findings of the trial. The approach that will be taken to define 'true TB status' should be described in greater detail. I suggest, as a minimum, that an attempt be made to obtain a baseline sputum sample for TB culture in all those starting TB treatment, regardless of study arm.

Is the rationale for, and objectives of, the study clearly described?
Partly

Is the study design appropriate for the research question?
Partly

Are sufficient details of the methods provided to allow replication by others?
Partly

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Epidemiology, TB transmission, infection control

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 23 Jul 2018
Peter MacPherson, Liverpool School of Tropical Medicine, Liverpool, UK

We are grateful for the detailed and helpful reviews provided by Tom Yates. We have responded in detail to all comments below.

1. I share Frank's concerns about the potential for the trial interventions to alter clinical practice in the standard of care arm. The authors acknowledge this possibility and state that, should the intervention appear effective in this trial, they might proceed to a cluster RCT, randomising clinics. I would note that the Hawthorne Effect here would be more likely to bias the effect estimate towards the null, rather than lead to a spurious positive result.

Thank you – we considered a number of different study designs, including a clinic cluster-randomised trial (would require an extremely large budget and massively increased resources), to an observational before-and-after design (with concomitant greater concerns around bias). No study design is perfect, but we hope that the single clinic pragmatic trial design we selected will allow us to provide strong, policy-relevant evidence. Should interventions be cost-effective in this trial, there would be a compelling argument for a subsequent cluster-randomised trial, or implementation evaluation as part of scale-up. We believe that any Hawthorne effect wold be more likely to bias effect estimate towards the null.

2. I think Frank's comment about the primary outcome measure not capturing the number of people initiating TB treatment is an important one. I would add that, ideally, the primary outcome measure should consider whether initiation of TB treatment was appropriate or not. Empirical TB treatment may be more common in the standard of care arm, where clinicians may have less ready access to investigations. Early initiation of TB treatment in people who do not have TB is not a good thing. I would like to see more discussion about this possibility.

We agree this is a difficult area. Right now, rates of empirical treatment are high, but exit interviews and high rate of prevalent TB (particularly among ART initiators) suggest that still far too few people with TB symptoms are evaluated appropriately, diagnosed, and linked promptly to treatment. We suspect that there is substantial under-diagnosis of TB among clinic attenders with current symptom-screen/sputum-based diagnostic approaches. We agree that clearly understanding the proportion of TB patient who are correctly started on TB would be of benefit. However, we believe that there is potentially greater public health benefit to be obtained by optimising available screening tests and linkage systems, and (hopefully) reducing time to treatment. We will be able to evaluate this to some extent by examining the proportion of participants with undiagnosed microbiologically-confirmed TB at Day 56.

Mathematical modelling studies suggest that an algorithm comprised of an initial chest x-ray, followed by confirmatory GeneXpert is likely to have achieve high sensitivity and specificity at acceptable programme costs. As discussed in detail in response to Point 3 below, the considerable additional resources required to ascertain microbiological TB status among all participants, coupled with the potential to compromise the study validity and strength of the evidence available from the trial, means that this approach is not feasible.

3. Not having cultures from all participants at baseline will make interpretation of the results difficult, particularly if there is a lot of empirical TB treatment (as I expect there may be). A good number of these patients may not have TB and unpicking that at day 56, by which time true cases of TB may have culture converted, may not be possible. Possible approaches to dealing with this issue...

a) Take sputum for culture from everyone at baseline but don't feed the result back to clinicians until day 56. Is this morally less acceptable than deliberately not obtaining a sputum specimen? I would point out that delaying sputum collection until day 56 would lead to treatment initiation being further delayed in patients whose smear negative/Xpert negative TB had not already been diagnosed.

b) Request that a baseline sputum sample for culture be taken from everyone in whom a decision to initiate TB treatment is made.

Thank you for raising these important issues. The trial team considered this in some detail. Given that TB culture is not routinely available in Malawi and most other countries in sub-Saharan Africa, we felt that doing sputum culture for TB for all participants at baseline and providing results to participants would not reflect current standard of care, and would provide evidence that was of lesser direct relevance to policymakers. We note that in the TBNeat Trial (Theron et al, Lancet 2014), although rates of empirical TB treatment were high, study TB culture taken at baseline and reported back to clinic nurses did provide additionality in TB treatment initiations; in the Smear Microscopy and Xpert Arms respectively, 18/154 (12%) and 15/170 (9%) of TB treatment initiations were attributed to a positive TB culture result. In Malawi under current standard of care, these patients would not achieve a microbiologically-confirmed diagnosis. It is possible that a baseline culture in our setting might attenuate differences in TB treatment initiations between groups, whilst diminishing the relevance of evidence we are able to provide to policymakers in the region.

We didn't feel it would be ethically acceptable to collect sputum and perform TB culture, but not report results to patients or clinicians, especially given the precedence already set in studies such as Theron et al.

We did carefully consider doing TB cultures for participants initiating TB treatment. However, there are a number of constraints to this. TB treatment will not be initiated by study Research Assistants, but by routine facility staff. In Groups 1 and 2, Study Research Assistants will not be present when TB treatment is initiated; all decisions to initiate TB treatment and registration procedures will be done by routine clinic healthworkers. In Group 3, participants may be linked to TB registration by study Research Assistants; however, it is possible that healthworkers will make a clinical decision to initiate TB treatment between Day 0 and Day 56 for participants not linked to registration by study team members at baseline. In such a case, there would be no Research Assistants present to collect sputum. Additionally, in all three groups, participants may initiate TB treatment at another clinic or hospital in Blantyre during the follow-up period (for example, they may be admitted to the central hospital, and registered for TB treatment there by routine healthworkers).

As part of a separate cluster-randomised trial (the HitTB Study), we attempted to collect a single spot sputum sample from all patients registering for TB treatment at all 18 health facilities in Blantyre. Citywide, we successfully collect sputum from 73% of all registering TB cases, demonstrating the challenges of achieving high levels of case ascertainment under routine programmatic conditions.

Overall, we remain unconvinced that attempting to collect a sputum sample for culture from all participants registering for TB treatment would be feasible, affordable, or sufficiently free of bias. Indeed, we suspect that there would be likely substantially greater ascertainment of TB culture status among participants in Group 3 than from the other two groups, hindering our ability to draw meaningful conclusions.

4. It would be good to see much more detail about what will happen during the day 56 study visit. If systematic recording of the adverse outcomes is to be achieved (especially 'Misclassification or misinterpretation of results leading to a participant starting TB therapy in error'), then presumably some determinantion will need to be made about which patients truly have TB and which do not. More detail about how this will be done should be included in the protocol, including how it will be done in patients that die before day 56.

We described the procedures undertaken at Day 56 as follows:

"Following completion of trial interventions, all participants in each of the three groups will be given a written appointment card to attend a follow-up assessment at the study research clinic room 56 days after randomisation. They will also be issued with a voucher that they can use to reimburse the cost of transport to the clinic for this assessment. Participants who don't attend their day 56 appointment will be traced to home.

To evaluate the primary trial outcome (time to tuberculosis treatment initiation), we will undertake a detailed questionnaire to record the date of TB treatment initiation, and verify by inspecting participant-carried TB treatment cards, medication and clinic TB treatment registers.

To evaluate the prevalence of undiagnosed tuberculosis, we will collect sputa from all participants. Samples will be transported to the TB Research Laboratory at the College of Medicine of Malawi, where they will be cultured for tuberculosis using the MGIT system, undergo smear examination using fluorescence microscopy, and tested using the GeneXpert MTB/Rif assay. Positive TB results will be reported to participants within three days of receipt (including by home tracing), and participants will be supported to register for TB treatment at the health facility.

To evaluate the prevalence of undiagnosed HIV at Day 56, we will offer all participants HIV testing, unless they are confirmed to be taking antiretroviral therapy. All participants requiring additional care will be supported to access either the HIV clinic, TB clinic, or outpatient clinic at the study clinic as required. Additionally, we will support referral and access to Queen Elizabeth Central Hospital should further specialist care be required.

To evaluate health-related quality of life (HRQoL), we will use the EQ-5D-3L. The EQ-5D is a generic

HRQoL measure, and was translated into Chichewa following international and EuroQoL guidelines. The EQ-5D-3L tool will be administered to all trial participants at baseline, and on Day 56."

We respectfully disagree that outcome assessment requires a definitive classification of whether a participant truly had TB or not. As discussed in response to Point 3 above, our trial primary outcome sets out to investigate time to TB treatment initiation among all participants, reflecting the reality of how TB care is currently delivered in low-resource settings. Secondary outcomes will evaluate the prevalence of undiagnosed TB, and proportion of participants achieving same-day treatment of TB. We don't think it is feasible to obtain a definitive microbiological TB status on all participants, as this would compromise our ability to provide unbiased evidence that is comparable to how diagnosis and care are currently delivered in Malawi and other countries in the region.

Relating to the adverse outcome of misclassification of results, we will record and report instances where a study-conducted GeneXpert is mistakenly reported to a participant as positive, when in fact results indicate it was negative (or vice-versa); or where a study-conducted culture is either mistakenly reported to a participant as positive, or whether misinterpretation leads to a delay in reporting results to a participant.

All participants who don't attend their day 56 outcome assessment will be traced to home, and their household members interviewed to ascertain TB and HIV treatment status in the period prior to death. We will not be able to take sputum samples from participants who have died before Day 56, so these secondary outcomes will be missing. In the preplanned sensitivity analysis (see the Statistical Analysis Plan in the Supplemental Material) we describe how this will be addressed:

"Participants who have missing information for outcomes will be excluded from primary analysis. However, in sensitivity analysis, we will use multiple imputation by chained equations to replace missing outcome variables."

5. It seems likely that clinicians making this assessment will need to speak to the patient (perhaps their relatives or clinicians who looked after them in hospital where patients have died) and to consult clinic notes from the first study visit. I struggle to see how this can be done whilst blinding the clinician to the group to which the patient was randomised, as is suggested will happen in the protocol.

The protocol describes procedures for ascertaining TB treatment initiation status at the Day 56 outcome visit:

"To evaluate the primary trial outcome (time to tuberculosis treatment initiation), Research Assistants will undertake a detailed questionnaire to record the date of TB treatment initiation, and verify by inspecting participant-carried TB treatment cards, medication and clinic TB treatment registers."

No clinicians will undertake outcome assessments; these will be done by study Research Assistants. In the Outcome Assessment section, we have replaced the word "we" with

"Research Assistants" to add clarity. All participants (regardless of group) will receive the same procedures at Day 56 assessment, meaning that the likelihood of differential outcome assessment is reduced. We will only interview caregivers during home tracing for participants who have died.

6. I would suggest tempering the comments about this study being pragmatic. It strikes me that the trial is somewhere on the continuum between an efficacy trial and a truly pragmatic trial. Were this a truly pragmatic trial, the trial interventions would be implemented by existing clinic staff. Here, study staff implement the interventions. I would like the authors to comment on how they will unpick the impact of the trial interventions from the impact of the additional human resources available to patients randomised to Groups 2 and 3. It seems possible (even likely) that the trial interventions would be less effective if they were being implemented by busy clinic nurses?

We agree that the definition of a pragmatic trial exists on a spectrum. Having interventions delivered by routine clinic staff is not the sole criteria that defines a pragmatic trial. Important other considerations include: whether a single intervention, or multiple complex interventions are evaluated; whether patient-focused/important measures (as opposed to biological parameters) are used as outcomes; whether routine health system data are used to assess outcomes; whether interventions are integrated within routine healthcare delivery. These issues and definitions are reviewed in Sox et al JAMA 2016. In this study, interventions (TB testing, HIV testing, linkage to treatment, other care services) may be also be provided through the routine clinic system by routine healthworkers in all three trial Groups, as shown in Figure 3. We feel that the design of the study places it firmly towards the pragmatic end of the spectrum.

Note, when we specify "clinicians" we mean non-physician healthworker cadres (Clinical Officers, Nurses), who provide much of the care in primary health care clinics. In the study clinic, there are no physicians. This reflects how care is delivered in primary clinics in much of sub-Saharan Africa. We hope that by conducting the study in a typical primary health care clinic, we will be able to provide evidence for cost-effectiveness under "real-life" conditions.

7. I am not a health economist but wonder whether any differences in EQ5D scores might be diluted by including in this comparison lots of people who don't have TB (in whom the impact of the intervention may be minimal).

We expect that, if random allocation is successful, the proportion of participants with TB in each arm will be similar. This outcome will evaluate the effect of interventions on change in EQ5D score among adults attending primary care with symptoms of TB. The Reviewer is correct that there is likely to be a wider range of EQ5D scores at baseline than if we included only patients with "true" TB, however, we don't think this would be possible for reasons described above in response to Point 3. We believe, and ours and others studies show, that adults attending primary health care with symptoms of TB often experience poor quality of care, and only a small fraction are investigated appropriately for both TB and HIV. Here, we hypothesise that optimised screening and linkage to care interventions for TB and HIV could improve the quality of life for adults with symptoms of TB.

8. Frank's comment about non TB findings on chest x-ray is an important one. Might this trial be a good opportunity to quantify the burden of non TB disease that might be missed and whether that varies depending on whether clinicians have access to CAD?

This is a good suggestion for additional exploratory analysis. Thank you.

9.With regards the second secondary outcome, I am unclear as to the rationale for excluding people that have stopped TB treatment for good reason (e.g. hepatotoxicity)?

Thanks – this was purely a pragmatic decision. Usually it is not possible to determine reason for TB treatment being stopped from TB treatment cards or registers. Therefore, we took the decision to exclude all such cases from the numerator.

10. With regards the nested evaluation of the CAD4TB tool, gold standard definitions of TB have been defined. Do gold standard definitions of 'not TB' need to also be defined, to allow estimates of specificity? I am not clear what is meant by 'reader characteristics' in the final paragraph of this section.

Thanks – We have defined a gold-standard for bacteriologically-confirmed TB as:

  - *"Bacteriologically-confirmed pulmonary tuberculosis, defined as either a documented positive GeneXpert result for Mycobacterium tuberculosison at least one sample of sputum taken for study purposes at Day 56 assessment; or documented growth of Mycobacterium tuberculosisand positive speciation using MPT 64 antigen tests on at least one culture of sputum taken for study purposes at Day 56 assessment; or documented identification of acid fast bacilli on at least one sputum sample taken for study purposes and examined by sputum smear microscopy at Day 56 assessment."*

Then the definition of "not bacteriologically-confirmed" is therefore the inverse of this.

"Reader characteristics" relate to whether radiologists are UK- or Malawi-based, and their years of experience. We have added to the protocol to clarify this.

11. People with previous TB make up a significant proportion of those with TB disease in high burden settings (Florian Marx has done some nice work describing this phenomenon). These people may have residual x-ray changes and there are reports of false positive Xpert tests in individuals who have been recently treated. Will excluding people that have had TB within the preceding six months deal with the latter issue, or can DNA persist for longer than this in people with recent successfully treated TB?

Thanks – yes, this was the reason for having this exclusion criteria. We agree that it is not perfect, but unfortunately no perfect TB diagnostic currently exists. Through this study, we hope that we can provide evidence for "optimised" (i.e. substantially better than current) diagnostic and treatment linkage approaches.

12. I would like some justification for the Bayesian analysis to be included in the protocol. What is this achieving that couldn't be achieved using a frequentist approach?

There is extensive literature describing the potential advantages (and potential complementarity) of Bayesian trial analysis (e.g., for introductory text see: *Bayesian Approaches to Clinical Trials and Health-Care Evaluation* by Speiglehalter, Abrams and Myles, Wiley and Sons, London 2004).

Some particular advantages of the Bayesian trial analysis include:

○ Explicit focus on how the trial should change our pre-held conceptions of how effective the trial interventions are (i.e. "a formalisation of learning from experience"). This includes the concept of "credibility of 'statistically significant' results once prior belief distributions are incorporated). We think this will be important in aiding decision-makers considering whether interventions should be implemented, and will mitigate against some of the issues raised by the reviewers.
○ Given that we are not aware of previous trials of these intervention strategies, we elected to elicit prior beliefs from key stakeholder groups. Prior beliefs of key stakeholder groups important in health decision-making (researchers, policymakers, clinicians, patients, community-members) can be incorporated into analysis, leading to explicit acknowledgement of constraints to acceptance.
○ Posterior distributions intervals from Bayesian analysis are interpretable in a manner than many researchers mistakenly think confidence intervals from frequentist analysis are.

13. Comments on the trial protocol
   ○ **It seems like enrollment will precede randomisation but I would like to see a more explicit comment on allocation concealment.**

In the protocol, we now state:

"Enrolment and baseline questionnaires will precede randomization. Randomisation will be done by research assistants using a random number allocation schedule running on study data-collection electronic tablets. Because of the nature of the study and the interventions offered, it will not be possible to blind participants or research assistants to allocation groups. Nevertheless, extensive steps will be taken to ensure that research assistants undertaking outcome assessments are blinded to participants' group. Additionally, the investigators, including the chief investigator and trial statistician, will remain blinded to allocation groups until final analysis. No unblinded interim analysis will be conducted."

14. I would like to see more comment about how patients who transfer their care out or are lost to follow up will be handled in the analysis.

This is described in detail in the Statistical Analysis Plan (Supplemental Material).

"Missing data will be examined for each variable and for each individual participant. A systematic assessment of missingness will be conducted to ascertain the reason and possible mechanism for missing data by identifying the quantity of missing data and patterns within the data. Missingness will be compared between randomised arms to assess for systematic biases.

Participants who have missing information for outcomes will be excluded from primary analysis. However, in sensitivity analysis, we will use multiple imputation by chained equations to replace missing outcome variables."

15. The CAD threshold that will be used should be pre-specified in the protocol.

Thank you. Yes, once the threshold has been established through evaluation of pilot films (to be done with the CAD4TB team in Blantyre in September), we will add this to the protocol.

16. A note should be made on the approach that will be taken to x-rays in women of childbearing age.

In this study, we follow recommendations in the WHO Guidelines for Chest Radiography in Tuberculosis Diagnosis, which state:

"For pregnant women and the fetus, a CXR does not pose any significant risk, provided that good practices are observed, as the primary beam is targeted away from the pelvis"

We have recruited experienced radiologists who have received additional training from expert radiologists and x-ray equipment manufacturers, and have procured lead shielding screens to protect the pelvis.

17. A note should be made about the approach that will be taken to patients that have RpoB mutations detected by Xpert.
Rifampicin resistance is rare in Blantyre and Malawi and we anticipate detecting very few to no cases. We have added text to state that:

"Participants with Rifampicin resistance detected on GeneXpert will be traced and supported to access the TB clinic at the Queen Elizabeth Central Hospital for further clinical assessment and evaluation for treatment."

18. It seems unlikely that the trial team will be able to see all participants on day 56. I suggest that the team allow themselves a window of time in which to undertake these study visits.

Thanks yes – we have added additional text to make this clear.

19. Please include an explicit comment about whether/how the urine LAM result will

be fed back to clinicians.

In this section, we refer to undertaking the nested diagnostic accuracy study of a novel manufacturer version of the urine LAM kit (FIND/Fujifilm), which has not yet undergone approval processes. As the performance of this assay is unknown, we will not report these results to clinicians.

20. There are a number of grammatical errors and the writing, in places, could be clearer. The section on trial outcomes, in particular, could be better written.

Thank you, we have carefully reviewed the manuscript and corrected errors.

Conclusions

21.The proposed two step approach to evaluating people with suspected pulmonary TB in settings such as these seems to be a good idea. It is good to see this proposal to evaluate the approach in a randomised controlled trial. My main concerns about the protocol concern the choice of primary outcome, particularly how this might perform in a setting where clinicians often initiate empirical TB treatment. An attempt to ascertain 'true TB status' would be valuable in interpreting the findings of the trial. The approach that will be taken to define 'true TB status' should be described in greater detail. I suggest, as a minimum, that an attempt be made to obtain a baseline sputum sample for TB culture in all those starting TB treatment, regardless of study arm.

We thank the Reviewer for these helpful suggestions. As discussed above, we don't aim to define true TB status in all participants in this study, but rather evaluate the cost-effectiveness of interventions that may promote more rapid initiation of TB treatment, and reduce the prevalence of undiagnosed TB and HIV at 8-weeks. Collection of sputum for culture from registering TB cases is not without difficulties, especially where TB treatment will not be initiated by study staff, or may be done at sites other than the study clinic (for example if a participant is admitted to the central hospital, and initiated on TB treatment there).

We think that the considerable additional cost of additional sputum culture (which is not currently funded), and considerable difficulties in obtaining sputum samples from participants registering for TB treatment under routine conditions, makes this problematic within an already complex study. We suspect that, with these constraints, we would likely have differential ascertainment of culture status at treatment initiation between groups.

Competing Interests: No competing interests were disclosed.

Reviewer Report 04 June 2018

https://doi.org/10.21956/wellcomeopenres.15894.r33178

? **Frank Cobelens**
[1] Department of Global Health, Academic Medical Center, Amsterdam, The Netherlands
[2] Amsterdam Institute for Global Health and Development, Amsterdam, The Netherlands

This article describes the rationale, objectives and protocol for a pragmatic, individually randomized 3-arm trial of the effectiveness and cost-effectiveness of optimised HIV and TB diagnosis and linkage to care interventions among adult primary care clinic attendants in Malawi. The trial addresses the question whether HIV self-testing and prompt linkage to HIV care for adults with TB symptoms and a TB triage testing algorithm based on digital chest x-ray with computer-aided diagnosis and Xpert can improve time to treatment initiation. Both are timely and important. In many settings in Africa HIV testing rates remain low, even in patients with presumptive TB, and many patients presenting with cough are not tested for TB despite availability of highly sensitive diagnostics such as Xpert. Triaging based X-ray is a promising approach to increase TB testing levels and case detection, and may have economic benefits over a policy (hardly implemented) of testing all patients meeting the criteria of presumptive TB.

Is the rationale for, and objectives of, the study clearly described?
Yes.

Is the study design appropriate for the research question?
The approach taken in this study is adequate although some concerns can be raised.

Design:
- The individually randomized design in a single clinic may introduce a considerable Hawthorne effect, by which routine practice (for HIV and TB testing in arm 1, and for TB testing in arm 2) may much improve over what is expected based on current practice. Although the authors acknowledge this in the discussion, they could have provided more details about how this risk will be mitigated and what the consequences might be in terms of sample size and statistical power.
- Both intervention arms comprise complex interventions in which several elements may be decisive for failure or success. Have authors thought about how these effects could be disentangled? Similarly, it would be interesting to read what safeguards are in place to optimize intervention fidelity.
- What is the currently the proportion of patients started on empirical TB treatment in this setting? As observed in various trials empirical treatment may obscure the effect of introducing Xpert on treatment initiation and outcomes including mortality, and also drive cost-effectiveness estimates. The subgroup analysis stratifying patients into either or not bacteriologically confirmed will shed light on this but the primary endpoint may well be affected.

Eligibility: the added value of a triaging approach is mainly in allowing more TB patients to be diagnosed earlier with limited added cost: if more patients reporting at primary care clinics (i.e. with less apparent TB symptoms) would be tested for TB by Xpert this may have major cost or logistics consequences. Triaging would allow identifying patients most at risk for having TB, which

will limit the number of Xpert tests that need to be done. However, in order to observe that benefit, eligibility should be sufficiently broad: not only patients who have cough as the presenting symptom, but potentially also those who report a cough regardless of the presenting complaint. It is unclear to what extent this is envisaged. It would also be informative to see how e.g. capacity limitations will be dealt with. If Xpert capacity is e.g. only 20 a day but 40 eligible patients report, how will the 20 be selected? Clearly this should not be based on clinical suspicion of TB.

Primary endpoint: time to treatment initiation seems a meaningful patient-important primary endpoint but it does not take into account the numbers of patients started on treatment. If for example in arm 1 three patients are started on TB treatment, all within 2 days, and in arm 300 with a mean of 2 days, the trial will show failure of the intervention. Unless one would count undetected cases at day 56 as started after day 56, but that does not seem the case.

Adverse events: one concern that has been raised about the use of CAD is that the algorithms were trained on TB patients, and that other pathology therefore may be missed. Should such occurrences be noted as well?

Are sufficient details of the methods provided to allow replication by others?
Yes. A few details could be described more clearly:

Eligibility: see above.

Cost-effectiveness: the interpretation of incremental cost-effectiveness ratios (ICERs) for triage strategies is not straightforward. For example the effectiveness in arm 1 or 2 could be higher than in arm 3 if all eligible patients in arm 1 or 2 would have Xpert testing, but with lower cost. It would be useful to see more detail about how ICERs for the triaging interventions will be interpreted.

Outcome evaluation: the active follow-up at day 56 is mentioned only late in the article while it is rather vital for understanding the endpoints. I would start with a broad design description that includes this element. Also no mention is made here of standard X-ray at day 56 whereas later paragraphs suggest that it this will be done for the CAD analysis.

Are the datasets clearly presented in a useable and accessible format?
It appears they will be.

Additional comments
- ○ The text has a few grammatical errors that could be addressed in a final version.
- ○ The discussion section states that "a triage approach (...) comprises an efficient, high sensitivity initial test, followed by a high specificity confirmatory test". This does not make sense to me. A test with both high sensitivity and high specificity would be a stand-alone diagnostic. The idea behind triaging is that the test must have high sensitivity but can have moderate specificity (a trade-off with monetary and non-monetary cost).

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?

Partly

Are sufficient details of the methods provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Yes

Competing Interests: No competing interests were disclosed.

I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard, however I have significant reservations, as outlined above.

Author Response 23 Jul 2018

Peter MacPherson, Liverpool School of Tropical Medicine, Liverpool, UK

We are grateful for the detailed and helpful reviews provided by Frank Cobelens. We have responded in detail to comments below.

1. The individually randomized design in a single clinic may introduce a considerable Hawthorne effect, by which routine practice (for HIV and TB testing in arm 1, and for TB testing in arm 2) may much improve over what is expected based on current practice. Although the authors acknowledge this in the discussion, they could have provided more details about how this risk will be mitigated and what the consequences might be in terms of sample size and statistical power.

Thank you for raising this issue. We agree that it is possible that a Hawthorne effect may occur. We carefully considered and ruled out a number of alterative study designs to mitigate against this. One approach would have been to do a clinic cluster-randomised trial; however, with three intervention arms this would be a massive undertaking, and require a substantial additional amount of funding and resources. We additionally considered a before-and-after implementation study, but were concerned that the quality of evidence provided would be substantially lower than that obtained with the current design.

Overall, we believe that the pragmatic study design we selected will provide the highest possible quality evidence to policymakers and researchers with the resources available. If this study demonstrates that interventions are cost-effective, we believe there would a compelling case to evaluate interventions in a larger cluster randomised trial, or as part of evaluation of intervention scale up. On the contrary, if interventions are not effective, this would suggest that further evaluation using these models may not be appropriate.

We have no data to estimate to what extent health workers might modify their practice whilst the study is running, and so are reluctant to speculate about impact on statistical power. However, we have included within the discussion section of the manuscript an additional sentence noting that the possibility of a Hawthorne effect might reduce the size of the differences between the standard of care arm and the interventions arms, and thus

power might be reduced.

Of note, one of the key strengths of the additional Bayesian analysis is that it will allow us to explore the posterior probability of effectiveness of interventions given the (potentially sceptical, neutral, or enthusiastic) prior beliefs elicited from a range of stakeholders (including researchers, clinic staff, and community leaders), allowing us to go some way towards addressing these issues.

2. Both intervention arms comprise complex interventions in which several elements may be decisive for failure or success. Have authors thought about how these effects could be disentangled? Similarly, it would be interesting to read what safeguards are in place to optimize intervention fidelity.

Our study was carefully designed to ensure that the delivery of interventions would reflect implementation under real-life conditions in low-resource settings. In keeping with the pragmatic trial design, interventions are delivered as a package (e.g. HIV self-testing and subsequent supported linkage to treatment). We believe that it is important that the cost-effectiveness of interventions is evaluated against conditions under which they would be implemented by national HIV and TB programmes in low-resource settings. Although this means that we cannot estimate the effectiveness of any single component of the trial interventions we believe that the evidence provided will be of greater relevance to policymakers grappling with how best to implement new and available HIV and TB diagnostic services. In reality, it would be unattractive if policymakers considered, for example, effectiveness of implementation of only CAD, without considering the additional requirements for confirmatory TB testing and linkage to care. By including combined interventions, we explicitly set out to evaluate the effectiveness and cost-effectiveness of interventions as they would be delivered under real-life conditions. The health economics analysis will allow assessment of the major intervention component drivers of cost-effectiveness.

With regards to ensuring fidelity to interventions: Pragmatic randomised trials explicitly acknowledge that intervention fidelity is not always perfect (e.g. see Sox et al, JAMA 2016). As such they allow estimation of effect sizes that are more realistically representative of implementation under real-life conditions. Nevertheless, as described in the publication and study protocol, we have taken a number of steps to optimise fidelity to interventions. These include: bio-identification of participants (by Simprints fingerprint scanning) at recruitment, prior to randomisation and allocation, prior to delivery of all trial interventions, and at outcome assessment; and exit interviews (when leaving the clinic) conducted on a random sample of 5% of participants to monitor quality of intervention delivery and fidelity to interventions. We will additionally collect data on process indicators, such as the numbers of chest x-rays performed, numbers of HIV tests done, and numbers of sputum samples evaluated.

3. What is the currently the proportion of patients started on empirical TB treatment in this setting? As observed in various trials empirical treatment may obscure the effect of introducing Xpert on treatment initiation and outcomes including mortality, and also drive cost-effectiveness estimates. The subgroup analysis stratifying patients

into either or not bacteriologically confirmed will shed light on this but the primary endpoint may well be affected.

We agree that there is likely to be a substantial proportion of patients initiated onto TB treatment empirically, as seen in previous studies. In Blantyre, we estimate that around half of TB patients are started on treatment without microbiological-confirmation. This is likely partly due to epidemiology, but also partly due to limited availability of diagnostics. This situation is common in much of sub-Saharan Africa. We selected the pragmatic trial design specifically acknowledging these facts. By comparing the effectiveness of interventions on time to TB treatment against current standard of care (where rates of empirical treatment are high), we should be able to determine whether optimised HIV testing and TB diagnosis results in more rapid initiation of treatment than what is currently done. Additionally, by testing all participants by TB culture, smear and Xpert at day 56 outcome, we should be able to go some way towards answering the question around appropriateness of TB treatment. Importantly, evidence will be directly relevant to policymakers, as we are comparing effectiveness against the current standard of care in the region TB diagnostic and treatment linkage system.

4. Eligibility: the added value of a triaging approach is mainly in allowing more TB patients to be diagnosed earlier with limited added cost: if more patients reporting at primary care clinics (i.e. with less apparent TB symptoms) would be tested for TB by Xpert this may have major cost or logistics consequences. Triaging would allow identifying patients most at risk for having TB, which will limit the number of Xpert tests that need to be done. However, in order to observe that benefit, eligibility should be sufficiently broad: not only patients who have cough as the presenting symptom, but potentially also those who report a cough regardless of the presenting complaint. It is unclear to what extent this is envisaged. It would also be informative to see how e.g. capacity limitations will be dealt with. If Xpert capacity is e.g. only 20 a day but 40 eligible patients report, how will the 20 be selected? Clearly this should not be based on clinical suspicion of TB.

We agree with these helpful points. Indeed, all acute adult attendees at the study clinic will be screened for the presence of cough, regardless of the reason for attendance – see study screening and eligibility procedures and Figure 1. In terms of capacity, we acknowledge that we may be limited by the availability of onsite GeneXpert testing. In terms of study conduct, we propose that we recruit consecutive daily attendee up to a limit depending upon the capacity of the study team and availability of Xpert, with details to be finalised pending completion of pilot work. We have added text to the paper to describe this.

"Where the number of eligible clinic attenders exceeds recruitment capacity, we will recruit participants up to a daily limit, with details finalized pending completion of pilot work showing the number of eligible participants per day."

In terms of any potential implementation following the study, we agree that this could be a limiting step. In our health economics analysis, we will model the additional costs and incremental cost effectiveness that would be achieved by National TB Programmes with a broader symptom screening initiative that would subsequently require greater numbers of

Xpert tests to be done.

5. Primary endpoint: time to treatment initiation seems a meaningful patient-important primary endpoint but it does not take into account the numbers of patients started on treatment. If for example in arm 1 three patients are started on TB treatment, all within 2 days, and in arm 300 with a mean of 2 days, the trial will show failure of the intervention. Unless one would count undetected cases at day 56 as started after day 56, but that does not seem the case.

Our previous data (Nliwasa et al, IJTLD 2016) showed that 17% of adults attending the clinic with TB symptoms initiated TB treatment within eight weeks. Since this outcome will be analysis as time-to-event, participants who do not initiate TB treatment will still contribute to analysis, being censored (without experiencing the primary trial outcome) at day 56. As a secondary outcome, we will compare between groups the prevalence of undiagnosed TB (by culture, smear and Xpert) among all participants (regardless of TB treatment status). Although some participants started on TB treatment will have reverted to culture negativity, we hope that this will mitigate against this.

6. Adverse events: one concern that has been raised about the use of CAD is that the algorithms were trained on TB patients, and that other pathology therefore may be missed. Should such occurrences be noted as well?

Thank you – yes. All chest x-rays (in Group 3 at baseline, and in all participants at Day 56) will be clinically reported by Consultant Radiologists, with results reported to patients (by tracing if necessary) and healthworkers where another abnormality is detected.

7. Cost-effectiveness: the interpretation of incremental cost-effectiveness ratios (ICERs) for triage strategies is not straightforward. For example the effectiveness in arm 1 or 2 could be higher than in arm 3 if all eligible patients in arm 1 or 2 would have Xpert testing, but with lower cost. It would be useful to see more detail about how ICERs for the triaging interventions will be interpreted.

We thank the reviewer for raising the issue of how the ICERs for the interventions will be interpreted. We have added to the text to provide more detail of the within-trial cost-effectiveness analysis and how the ICER will be estimated and interpreted. In the example raised by the reviewer where Group 1 or 2 is more effective than Group 3, and associated with lower cost, the interpretation would be that Group 3 would be dominated as it represents a costlier and less effective alternative, and therefore would not be a cost-effective option.

8. Outcome evaluation: the active follow-up at day 56 is mentioned only late in the article while it is rather vital for understanding the endpoints. I would start with a broad design description that includes this element. Also no mention is made here of standard X-ray at day 56 whereas later paragraphs suggest that it this will be done for the CAD analysis.

Thank you. We have added text to the manuscript to give greater prominence to the Day 56

outcome assessment

8. The text has a few grammatical errors that could be addressed in a final version.

Thank you, we have carefully reviewed the text and corrected errors.

10. The discussion section states that "a triage approach (…) comprises an efficient, high sensitivity initial test, followed by a high specificity confirmatory test". This does not make sense to me. A test with both high sensitivity and high specificity would be a stand-alone diagnostic. The idea behind triaging is that the test must have high sensitivity but can have moderate specificity (a trade-off with monetary and non-monetary cost).

We agree. Here we refer to two separate tests used in sequence in this intervention group: Chest x-ray with CAD, followed by GeneXpert MTB/Rif if CAD indicates that TB is probable.

Competing Interests: No competing interests were disclosed.