



Cochrane
Library

Cochrane Database of Systematic Reviews

Symptom screening for active tuberculosis in pregnant women living with HIV (Protocol)

LaCourse SM, Cranmer LM, Bekker A, Steingart KR, Black D, Horne DJ, Oren E, Pals S, Modi S, Mathad J

LaCourse SM, Cranmer LM, Bekker A, Steingart KR, Black D, Horne DJ, Oren E, Pals S, Modi S, Mathad J.

Symptom screening for active tuberculosis in pregnant women living with HIV.

Cochrane Database of Systematic Reviews 2018, Issue 1. Art. No.: CD012934.

DOI: 10.1002/14651858.CD012934.

www.cochranelibrary.com

TABLE OF CONTENTS

HEADER	1
ABSTRACT	1
BACKGROUND	1
Figure 1.	4
OBJECTIVES	5
METHODS	5
ACKNOWLEDGEMENTS	8
REFERENCES	8
APPENDICES	11
CONTRIBUTIONS OF AUTHORS	20
DECLARATIONS OF INTEREST	20
SOURCES OF SUPPORT	20

[Diagnostic Test Accuracy Protocol]

Symptom screening for active tuberculosis in pregnant women living with HIV

Sylvia M LaCourse¹, Lisa M Cranmer², Adrie Bekker³, Karen R Steingart⁴, Danae Black⁵, David J Horne⁶, Eyal Oren⁷, Sherri Pals⁸, Surbhi Modi⁸, Jyoti Mathad⁹

¹Department of Medicine, Division of Allergy and Infectious Diseases, University of Washington, Seattle, Washington, USA.

²Department of Pediatrics, Emory University School of Medicine and Children's Healthcare of Atlanta, Atlanta, Georgia, USA.

³Department of Paediatrics, Stellenbosch University, Cape Town, South Africa. ⁴Cochrane Infectious Diseases Group, Liverpool School of Tropical Medicine, Liverpool, UK. ⁵Department of Epidemiology, University of Washington, Seattle, Washington, USA.

⁶Department of Medicine, Division of Pulmonary and Critical Care Medicine, and Firland Northwest TB Center, University of Washington, Seattle, WA, USA. ⁷Division of Epidemiology and Biostatistics, San Diego State University, San Diego, California, USA.

⁸Division of Global HIV and Tuberculosis, Center for Global Health, US Centers for Disease Control and Prevention (CDC), Atlanta, Georgia, USA. ⁹Department of Medicine, Division of Infectious Diseases, Center for Global Health, Weill Cornell Medical College, New York, USA

Contact address: Sylvia M LaCourse, Department of Medicine, Division of Allergy and Infectious Diseases, University of Washington, 325 9th Avenue, Box 359931, Seattle, Washington, WA 98104, USA. sylvial2@u.washington.edu.

Editorial group: Cochrane Infectious Diseases Group.

Publication status and date: New, published in Issue 1, 2018.

Citation: LaCourse SM, Cranmer LM, Bekker A, Steingart KR, Black D, Horne DJ, Oren E, Pals S, Modi S, Mathad J. Symptom screening for active tuberculosis in pregnant women living with HIV. *Cochrane Database of Systematic Reviews* 2018, Issue 1. Art. No.: CD012934. DOI: 10.1002/14651858.CD012934.

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

ABSTRACT

This is a protocol for a Cochrane Review (Diagnostic test accuracy). The objectives are as follows:

To assess the accuracy of the four-symptom screen (cough, fever, night sweats, or weight loss) for identifying active TB in pregnant PLHIV who are screened in an outpatient or community setting.

- To investigate potential sources of heterogeneity of the accuracy of the four-symptom screen between studies including: ART status, CD4 cell count, gestational age, pregnancy stage (pregnancy vs. postpartum), screening test definition of cough (any cough vs. cough greater than 2 weeks).
- To describe the accuracy of single symptoms included within the four-symptom screen, additional symptoms or symptom combinations, for identifying active TB in pregnant PLHIV. For example, additional symptoms may include failure to gain weight or fatigue.

BACKGROUND

Tuberculosis (TB) is the leading cause of mortality in people living

with HIV (PLHIV) ([UNAIDS 2016](#)); and the third leading cause of death among women of child-bearing age in high TB burden areas ([Mathad 2012](#)). Every year, over 200,000 pregnant women

Symptom screening for active tuberculosis in pregnant women living with HIV (Protocol)

Copyright © 2018 The Cochrane Collaboration. Published by John Wiley & Sons, Ltd.

1

are estimated to develop TB during pregnancy (Sugarman 2014), though this is likely an underestimation due to administrative, diagnostic, or possible immunologic-related delays in diagnosis (Llewelyn 2000; Zenner 2012). Earlier detection and treatment of TB in pregnant PLHIV is crucial in improving outcomes for both mothers and their young children (Getahun 2012). The World Health Organization (WHO) recommends that PLHIV, including pregnant women, should be routinely screened for symptoms of active TB at every health facility visit using a four-symptom screen (cough, fever, night sweats, and weight loss (WHO 2011a). Routine TB screening in settings of high HIV and TB burden offers opportunity to identify PLHIV who require further diagnostic work-up to evaluate for active TB, as well as PLHIV without symptoms who are unlikely to have TB and may be candidates for isoniazid preventive therapy (IPT) (WHO 2011a). Recent data suggest that the performance of the four-symptom screen for active TB may perform differently in pregnant PLHIV compared to PLHIV in general (Hoffmann 2013; Kancheya 2014; LaCourse 2016; Modi 2016). For this systematic review, we will review the accuracy of symptom screening in pregnant PLHIV for active TB.

Target condition being diagnosed

TB is an airborne infectious disease caused by organisms within the complex *Mycobacterium tuberculosis* (*Mtb*) (Pai 2016). Although typically a disease of the lungs, *Mtb* can disseminate throughout the body. Infection primarily occurs after close and prolonged exposure to a person with active TB, with *Mtb* spread often occurring through coughing. The majority of people infected with *Mtb* will not develop active TB and are often described as having latent TB infection (LTBI). They are largely asymptomatic and not infectious. Although previous descriptions have categorized LTBI and active TB as binary states, in reality there is likely a spectrum: from clearance of the infection with evidence of immune response, to LTBI and granuloma formation, to subclinical disease (typically without symptoms but potentially infectious), to active (symptomatic) disease (Lawn 2011). Active TB can occur after recent exposure and infection, or much later in the setting of progression of previously quiescent LTBI. Approximately 5% to 15% of people will develop active TB in the setting of recent infection (within a few months or years), with the remaining described as having LTBI who are then at risk of developing TB in the future (Pai 2016).

Progression to active TB is more likely to occur in the setting of immune suppression with HIV or malnutrition (Horsburgh 2011). HIV significantly increases the risk for progression from LTBI to active TB (Lawn 2011), from an approximately 10% lifetime risk to over 10% per year (Corbett 2003). Given this elevated risk of active TB, the WHO recommends that PLHIV in TB-endemic areas be provided IPT to reduce the risk of progression from LTBI to active TB (WHO 2011a). The risk of active TB also appears to be elevated in pregnant and postpartum women (Zenner 2012).

The time from LTBI to active TB (when it occurs) can be variable, and it can take months to years for individuals to develop symptoms and bacteriologically detectable TB. People with active TB who are not on treatment can continue to be infectious for prolonged periods of time. Untreated maternal active TB is associated with poor maternal and infant outcomes (Getahun 2012; Mathad 2012). TB diagnosis may be delayed in pregnant women in part due to masking of symptoms from physiologic (Hamadeh 1992) or immunologic changes (Kourtis 2014; Mathad 2012; Singh 2007) associated with pregnancy. Maternal HIV and TB co-disease is associated with adverse infant outcomes including prematurity, small for gestational age, vertical HIV transmission, neonatal TB, and death (Getahun 2012; Mathad 2012). Treatment for TB diagnosed during pregnancy should not be delayed (Nahid 2016; WHO 2010). In a recent large case-control study in South Africa, mothers with HIV and TB were more likely to experience poor maternal and infant outcomes despite widespread access to antiretroviral therapy (ART), TB treatment and infant IPT (Salazar-Austin 2017).

TB symptom screening for PLHIV

The WHO defines screening for TB as the “systematic identification of people with suspected active TB, in a predetermined target group, using tests, examinations or other procedures that can be applied rapidly” with the goal to “efficiently distinguish people with a high probability of having active TB from those who are unlikely to have active TB” (WHO 2013). Screening tests are not intended to be diagnostic, rather after a positive screen result, diagnosis is established using additional diagnostic tests. Systematic screening for TB is primarily provider-initiated, and allows for earlier detection of TB and permits earlier treatment and improved outcomes (WHO 2013). In contrast, “passive-case finding” often relies on people with potential TB reporting their symptoms to providers prior to further work-up.

As part of the guidelines for intensified TB case finding and isoniazid preventive therapy (IPT) for PLHIV in resource-limited settings, the WHO recommends that PLHIV should be routinely screened for active TB at every health facility visit using a four-symptom screen (cough, fever, night sweats, and weight loss: CFSW) (WHO 2011a). The presence of any one of these four symptoms is considered a positive screen. The absence of all four symptoms is considered a negative screen. The goal of TB symptom screening in PLHIV is two-fold: 1) to identify those with a positive symptom screen who should undergo further evaluation for active TB; and 2) to identify those with a negative symptom screen who are unlikely to have TB and therefore should be offered IPT. The recommendation regarding the use of the TB four-symptom screen in PLHIV is based primarily on the individual participant data meta-analysis performed by Getahun and colleagues of 12 observational studies including 8148 PLHIV (Getahun 2011). The focus of that meta-analysis was to identify a combination

of clinical symptoms that could be readily identified at any level of health system with the combination of greatest sensitivity and specificity, as well as negative predictive value, for active TB among PLHIV. The best performing rule included the presence of any one of CFSW with a sensitivity of 78.9% (95% CI 58.3% to 90.9%) and specificity of 49.6% (95% CI 29.2% to 70.1%). Negative predictive values were 99.6% (95% CI 99.5% to 99.6%) and 97.7% (95% CI 97.4% to 98.0%) for prevalences of 1% and 5%, respectively.

TB symptom screening for pregnant PLHIV

Integrating TB case-finding within maternal and child health settings is considered a fundamental part of the Global Plan to End TB 2016-2020 ([Stop TB Partnership 2015](#)). The WHO guidelines specifically include TB screening of pregnant PLHIV using the four-symptom screen ([WHO 2011a](#)). However, a number of recently published studies report that the four-symptom screen appears to have lower sensitivity among pregnant PLHIV ([Hoffmann 2013](#); [Kancheva 2014](#); [LaCourse 2016](#); [Modi 2016](#)), compared to PLHIV in general. Previously published estimates of the high negative predictive value in pregnant PLHIV ([Gupta 2011](#)), may be different in the current era of widespread ART ([Ahmad Khan 2014](#); [Rangaka 2012](#)), including ART initiated for prevention of mother-to-child transmission (PMTCT) of HIV in pregnant and breastfeeding women and continued for life regardless of CD4 count (referred to as “Option B+” by the WHO) ([WHO 2015a](#)).

Index test(s)

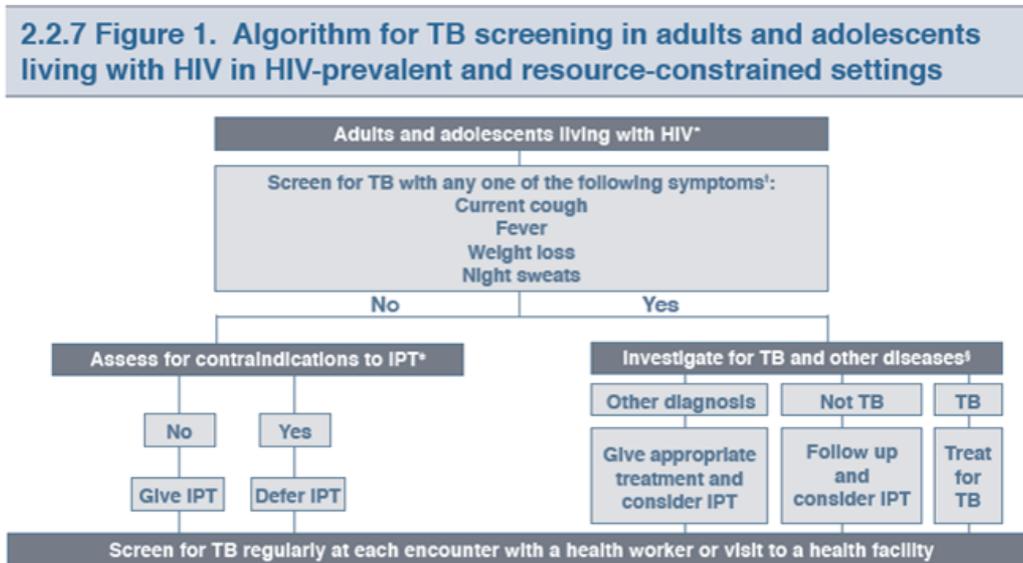
The primary index test that we will evaluate in this review is the WHO-recommended TB four-symptom screen, which consists of questions regarding the presence of any cough (of any duration), fever, night sweats, or weight loss (CFSW) to identify pregnant PLHIV with active TB. The four-symptom screen consists of “Yes or No” type questions regarding the presence of any CFSW symptoms. The presence of any one of these symptoms is considered a positive screen. The absence of all four of these symptoms is con-

sidered a negative screen. The four-symptom screen is designed to be easily asked by healthcare providers in an outpatient or community setting and can be done efficiently in low-resource settings, allowing for further diagnostics to be reserved for those individuals who are most likely to have TB.

Clinical pathway

The four-symptom screen is the initial test for PLHIV in a clinical pathway that includes subsequent confirmatory testing (Xpert MTB/RIF (Xpert), and in some cases chest radiograph or mycobacterial culture (solid or liquid)), to establish the diagnosis of active TB ([Figure 1](#)) ([WHO 2011a](#); [WHO 2013](#)). PLHIV with a negative four-symptom screen are considered to be unlikely to have TB, and are assessed for IPT eligibility to prevent active TB ([WHO 2011a](#); [WHO 2013](#)). PLHIV with a positive four-symptom screen continue through a testing pathway and undergo further diagnostic tests for TB and other diseases. The WHO currently recommends that PLHIV with a positive screen have an Xpert (rapid nucleic acid amplification test for *Mtb*), performed on expectorated sputum as a primary diagnostic test for TB ([WHO 2011b](#)). If Xpert positive, individuals should be started on multi-drug anti-tuberculosis treatment. In some settings, a culture is performed if Xpert positive to confirm TB diagnosis or to confirm whether the *Mtb* is sensitive to first-line anti-tuberculosis treatment. In most resource-limited settings, routine confirmatory culture is not performed unless the patient has failed first-line therapy or had a relapse. When available, chest radiography can also be used as an additional screening test to improve the pretest probability of the subsequent diagnostic test, and to reduce the number of people who need to undergo further diagnostic evaluation ([WHO 2013](#)). However in many settings, chest radiograph is unavailable ([Saito 2012](#)), therefore the WHO does not recommend that it be required in the evaluation of individuals suspected of TB. Additionally, clinicians may be reluctant to perform a chest radiograph in pregnancy, despite the low risk of this amount of radiation exposure to a pregnant woman or her fetus ([Ratnapalan 2004](#)).

Figure 1. Algorithm for TB screening in adults and adolescents living with HIV in HIV-prevalent and resource-constrained settings. Reprinted from: Guidelines for intensified tuberculosis case-finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. Copyright: WHO 2011a.



FOOTNOTES TO ALGORITHM FOR ADULTS

- * Every adult and adolescent should be evaluated for eligibility to receive ART. Infection control measures should be prioritized to reduce *M. tuberculosis* transmission in all settings that provide care.
- † Chest radiography can be done if available, but is not required to classify patients into TB and non-TB groups. In high HIV-prevalence settings with a high TB prevalence among people living with HIV (e.g. greater than 10%), strong consideration must be given to adding other sensitive investigations.
- * Contraindications include: active hepatitis (acute or chronic), regular and heavy alcohol consumption, and symptoms of peripheral neuropathy. Past history of TB and current pregnancy should not be contraindications for starting IPT. Although not a requirement for initiating IPT, TST may be done as a part of eligibility screening in some settings.
- † Investigations for TB should be done in accordance with existing national guidelines.

Role of index test(s)

The role of the four-symptom screen is a triage test as part of an algorithm to identify PLHIV who require further investigations for TB (those with positive symptom screens), as well as those who should be assessed for IPT (those with negative symptom screens) (Figure 1). Triage tests are used before diagnostic tests with only those patients with a positive result continuing through a testing pathway (Bossuyt 2006). Triage tests may be less accurate than existing diagnostics tests and are not meant to replace them, but may have additional advantages such as ease of use or low cost. A true positive (TP) on the four-symptom TB screen allows for referral for further diagnostics leading to early diagnosis and treatment of active TB. A true negative (TN) allows for potentially costly diagnostics to be avoided, and for evaluation of IPT for those truly without TB. The consequences of false positives (FP) are further diagnostics with which active TB would be ruled out. The consequences of false negatives (FN) are a possible delay in diagnosis and treatment for active TB as well as the potential for

the inappropriate initiation of IPT. Individuals with active TB who receive isoniazid as opposed to multi-drug anti-tuberculosis therapy could be at risk of developing drug-resistant TB. However rates of isoniazid resistance among PLHIV receiving IPT in TB prevention trials are similar to baseline rates of isoniazid resistance, and this risk is thought to be relatively low (WHO 2015c).

Alternative test(s)

The four-symptom screen is considered the standard of care for the initial TB screening test for PLHIV (WHO 2011a). Alternative approaches to screening include passive case-finding, alternate symptom screening, sputum microscopy, chest radiograph, and urine lateral flow lipoarabinomannan (LAM) assay (LF-LAM). Passive-case finding relies on identifying TB among people who actively seek care for TB (WHO 2013). This is effective in identifying people who are highly symptomatic, but is less effective in identifying people who may have less pronounced or protean

symptoms which may often be the case in PLHIV.

Alternate symptom screening includes the presence of chronic cough (for example longer than two weeks). Chronic cough has low sensitivity for TB among PLHIV, leading to missed cases and diagnostic delays (Getahun 2011).

Similarly, sputum microscopy has poor sensitivity in PLHIV, who often have paucibacillary disease. Chest radiographs can be negative or inconclusive in PLHIV, and require expert trained staff to read the images. Therefore the WHO does not recommend or require the use of sputum microscopy or chest radiograph for the initial screening test for TB in PLHIV (WHO 2011a).

LAM antigen is a lipopolysaccharide present in mycobacterial cell walls, and can be detected in the urine of people with active TB using urine LF-LAM (Lawn 2012). Urine-based testing has potential advantages over sputum-based testing due to ease of collection and lower infection control risks compared to sputum collection. However, due to low sensitivity and specificity (Shah 2016), the WHO strongly recommends against using urine LF-LAM for TB screening in PLHIV in general, except for those PLHIV with low CD4 or who are seriously ill (WHO 2015b). Portable digital chest radiographs with computer-assisted diagnosis, and non-sputum based biomarkers may have a role for initial screening in the future, but are not currently recommended in the initial screening or triage for TB in PLHIV (UNITAID 2015).

Rationale

The aim of this systematic review is to synthesize evidence regarding the accuracy of the four-symptom screen for active TB in pregnant PLHIV.

This subject is particularly relevant as identifying PLHIV with TB to initiate timely TB treatment, and identifying PLHIV who are safe to initiate IPT, are cornerstones of key strategies of the Global Plan to End TB 2016-2020 (Stop TB Partnership 2015). The Maternal-Child Tuberculosis Working Group of the International Union Against Tuberculosis and Lung Disease has identified the performance of TB screening tools in pregnant PLHIV as an important research gap (Modi 2017 [pers comm]).

This review will focus on pregnant PLHIV for a number of reasons. The risk of TB in women, especially those with HIV, appears higher during pregnancy and postpartum periods (Deluca 2009; Hoffmann 2013; Kancheya 2014; LaCourse 2016; Modi 2016). Maternal TB is associated with poor maternal and infant outcomes, particularly among HIV-infected women and their children (Getahun 2012; Mathad 2012; Salazar-Austin 2017). Timely detection of TB in peripartum PLHIV may reduce TB-associated morbidity and mortality. Antenatal and PMTCT settings provide opportunity for routine TB symptom screening in pregnant PLHIV (Deluca 2009; Getahun 2012).

OBJECTIVES

To assess the accuracy of the four-symptom screen (cough, fever, night sweats, or weight loss) for identifying active TB in pregnant PLHIV who are screened in an outpatient or community setting.

Secondary objectives

- To investigate potential sources of heterogeneity of the accuracy of the four-symptom screen between studies including: ART status, CD4 cell count, gestational age, pregnancy stage (pregnancy vs. postpartum), screening test definition of cough (any cough vs. cough greater than 2 weeks).
- To describe the accuracy of single symptoms included within the four-symptom screen, additional symptoms or symptom combinations, for identifying active TB in pregnant PLHIV. For example, additional symptoms may include failure to gain weight or fatigue.

METHODS

Criteria for considering studies for this review

Types of studies

We will include cross-sectional and cohort studies in which pregnant PLHIV are tested with both TB symptom screening as well as at least one of the reference standards (Xpert or culture). In addition, we will include diagnostic case control studies. Case control studies may overestimate sensitivity and specificity (Lijmer 1999); however, as we anticipate identifying few relevant studies, we will include them. We will also include randomized controlled trials with each arm as a separate study. In addition, we will include published TB prevalence surveys; however we will exclude participants with known prevalent TB (i.e. on anti-tuberculosis treatment at time of screening). Data from baseline measurement in longitudinal cohorts, as well as interventional trials in which persons with TB need to be excluded, will be eligible. For longitudinal studies where incident TB is identified after enrolment, data regarding symptom screening which occurs at the time of incident TB diagnosis will be eligible.

We will focus on studies that clearly describe TB symptom screening in pregnant PLHIV either in their methods or results. We will include studies that report data from which we can extract true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) for the four-symptom screen. Studies with zero TB cases will be eligible for inclusion for estimates of specificity. Studies in which not all participants undergo the reference standard will be eligible for inclusion, as long as it is clear which participants underwent the index (symptom screening) and reference

tests (Xpert or culture). Studies that screened PLHIV or pregnant women in general but included pregnant PLHIV will be eligible, as long as data can be extracted from the study specifically for pregnant PLHIV.

Participants

We will include pregnant PLHIV of all ages eligible for TB screening and not yet known to have TB at time of screening, who were screened in the outpatient or community setting. This may include pregnant PLHIV from all types of populations including the general population of PLHIV (included in mass case finding in TB prevalence surveys) as well as specifically targeted populations (screening of household members of known TB cases).

Index tests

For the primary analyses, we will include studies which use the four-symptom screen (cough, fever, night sweats, weight loss: CFSW). A positive index test is the presence of any one of the four symptoms. A negative index test is the absence of all four of these symptoms. For our secondary analyses, we will also include studies which include additional symptoms (e.g. fatigue, failure to gain weight), and symptom combinations for identifying active TB.

Target conditions

The target condition is active TB.

Reference standards

Our primary reference standards will be sputum mycobacterial culture (liquid or solid) and Xpert, where 'TB' is defined as a positive culture for *Mtb* or a positive Xpert result, and 'No TB' is defined as a negative culture for *Mtb* or negative Xpert result (or if both tests are performed, then both tests are negative). In general, mycobacterial culture is considered the reference standard by which other TB diagnostics are assessed. For the purposes of this systematic review, mycobacterial culture will include both liquid (automated reading of mycobacterial growth inhibitor tubes (MGIT)) and solid medium (Löwenstein-Jensen (LJ)). MGIT is thought to have increased sensitivity for identifying *Mtb*, but potentially lower specificity due to higher rates of contamination (Somoskovi 2000). Xpert is a WHO-endorsed nucleic acid amplification (NAAT) for the diagnosis of TB, and is the recommended primary diagnostic test for PLHIV in resource-limited settings (WHO 2011b). In a recent systematic review, Xpert has a pooled sensitivity of 89% (95% Credible Interval (CrI) 85% to 92%) and pooled specificity of 99% (95% CrI 98% to 99%) as an initial diagnostic test as measured against mycobacterial culture as the reference standard, with slightly reduced pooled sensitivity among PLHIV, 79% (95% CrI 70% to 86%) (Steingart 2014).

For the purpose of this review, participants with mycobacterial culture negative for *Mtb*, but positive for nontuberculous mycobacteria will also be considered to have 'No TB'.

Search methods for identification of studies

We will attempt to identify all relevant studies regardless of language or publication status (published, unpublished, in press, and in progress).

Electronic searches

We will search the following databases using the search terms and strategy described in Appendix 1: Cochrane Infectious Diseases Group Specialized Register; Central Register of Controlled Trials (CENTRAL), published in the Cochrane Library; MEDLINE (OVID); Embase (OVID); CINAHL (EBSCOHost); Science Citation Index-Expanded (Web of Science) and Scopus. We will search the WHO International Clinical Trials Registry Platform (ICTRP) (www.who.int/ictip/en/), and ClinicalTrials.gov (<https://clinicaltrials.gov/ct2/home>), for trials in progress, using 'tuberculosis' and 'symptom screening' as search terms.

Searching other resources

We will contact researchers and experts in the field to identify any additional eligible studies. We will also check the references of all included studies to identify additional studies.

Data collection and analysis

Selection of studies

Two review authors will independently screen all titles and abstracts to determine potentially eligible studies. We will then obtain the full-text articles of these potentially eligible studies and two review authors will independently assess whether they should be included based on pre-defined inclusion and exclusion criteria. We will resolve disagreements by discussion between the two review authors, as well as a third author if necessary. We will contact study authors for clarification of methods and other information if necessary. Reasons for exclusion will be recorded and summarized in a 'Characteristics of excluded studies' table. We will illustrate the study selection process in a PRISMA diagram.

Data extraction and management

We have drafted a data extraction form (Appendix 2) and will pilot the form with at least two included studies. We will finalize the form based on this pilot. Two review authors will independently

extract data from each study using the finalized form. Both review authors will discuss any inconsistencies to obtain consensus. Any disagreements will be resolved either through consensus or by a third review author if necessary. We will enter into a database (either Excel or Research Electronic Data Capture [REDCap]) (Harris 2009). Extracted data will be stored on password-protected computers, or in the case of REDCap (www.iths.org/investigators/services/bmi/redcap/), in a password protected online web application. Extracted data will be stored for review updates and we will seek Cochrane and CDC approval prior to update.

Data extraction will include the following characteristics.

- Authors, title, publication year, journal, email address of corresponding author.
- Whether author contacted, dates of contact, author response.
- Year(s) study conducted.
- Language of the publication, publication status.
- Country where study conducted.
- Reference standard(s): sputum Xpert, culture (solid or liquid), number of samples per individual.
- Whether participants unable to produce sputum included in study.
- Clinical setting (outpatient, community screening), participant selection.
- Study design (including direction of study data collection i.e. prospective, retrospective).
- Index test(s): four-symptom screen (CFSW), additional symptoms (cough > 2 weeks, failure to gain weight, fatigue, etc.).
 - Purpose of screening as described in the study.
 - Number after screening by study inclusion and exclusion criteria.
- Number included in analysis (include if available # in study, # pregnant PLHIV, # screened - # withdrew).
- Patient characteristics: age (range, mean (SD), median (IQR));
 - HIV characteristics: HIV status of participants (#, %), ART status (combination ART vs. ART for PMTCT alone), CD4 (range, mean (SD), median (IQR)), HIV viral load (VL) (% with undetectable VL, range, mean (SD), median (IQR)).
 - Pregnancy characteristics: gestational age (range, mean (SD), median (IQR)), postpartum age (range, mean (SD), median (IQR)), pregnancy vs. postpartum.
 - Patients with a previous history TB (#, %).
 - Details of outcomes: # of true positives (TP), true negatives (TN), false positives (FP), false negatives (FN); # participants with missing and/or unavailable results.

We will classify country income status as 'low and middle income' or 'high income', according to the World Bank List of Economies (World Bank 2016). Additional tables may be created for other symptoms or symptom combinations if they are reported. Review authors who are also authors of primary studies will recuse themselves from the screening and data extraction of their own studies

and an alternate reviewer will perform the screening and data extraction.

Assessment of methodological quality

We will assess the methodological quality of included studies using the Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) instrument (Whiting 2011), which we will tailor to this review. The QUADAS-2 consists of four domains: (1) patient selection; (2) index test(s); (3) reference standard(s); and (4) flow and timing. We will assess all domains for risk of bias and the first three domains for concerns regarding applicability. As recommended, we will first develop guidance on how to appraise each question and interpret this information. Then, one review author will pilot the tool with two of the included studies. Based on experience gained from the pilot, we will finalize the tool. Two review authors will independently complete QUADAS-2. We will resolve disagreements through discussion or, failing that, arbitration by a third review author. We will present the results of the quality assessment in the text, table, and graphs. The preliminary tool with signalling questions tailored to this review is in Appendix 3.

Statistical analysis and data synthesis

We will perform descriptive analyses of the characteristics of the included studies using Stata 14 (StataCorp 2015), and present key study characteristics in 'Characteristics of included studies' tables. We will use data reported in the two-by-two tables to calculate sensitivity and specificity estimates and 95% confidence intervals (CI) for the accuracy of the four-symptom screen from individual studies using Review Manager 5 (RevMan 5) (RevMan 2014). When possible, TB cases diagnosed by CXR or clinical suspicion alone will be excluded from the extracted data. We will present individual study results graphically by plotting the estimates of sensitivity and specificity (and their 95% CIs) in forest plots and receiver operating characteristic (ROC) space using RevMan 5. If data allow, we will include descriptive analyses of the performance of single symptoms, additional symptoms, or different symptom combinations other than the four-symptom screen.

We will fit a bivariate random-effects meta-analysis model (Chu 2010; Macaskill 2010; Reitsma 2005) to estimate the pooled sensitivity and specificity and corresponding 95% CIs using the `meqrlogit` and `metandi` commands in Stata (version 14). We have selected the bivariate model because the four-symptom TB screen uses a common positivity criteria or threshold, i.e. the presence of any one of the four symptoms is considered a positive result. The bivariate random-effects approach will allow us to calculate the pooled estimates of sensitivity and specificity while dealing with potential sources of variation caused by: (1) variability of sensitivity and specificity estimates within individual studies; (2) correlation between sensitivity and specificity across studies; and (3) variation in sensitivity and specificity between studies. For some

subgroups or screen definitions we may not be able to give meaningful summary estimates of sensitivity and specificity, therefore they will be evaluated using descriptive methods.

If we find that, in using hierarchical models, the analyses fail to converge due to a small number of studies or sparse data, we will consider simplifying the models into fixed-effect models by eliminating the random-effects parameters for sensitivity or specificity (Takwoingi 2017).

Investigations of heterogeneity

We will examine the forest plots and ROC plots through visual examination for heterogeneity. If the data allow, we will analyse potential determinants or sources of heterogeneity as covariates in the models using meta-regression (Macaskill 2010). We will include the following pre-specified categorical study-level covariates.

- ART status (on ART vs. not ART).
- Pregnancy stage (pregnant vs. postpartum).
- Screening test definition of cough (any cough vs. cough longer than two weeks).

Sensitivity analyses

If there are sufficient data, we will perform sensitivity analyses to explore the contribution of risk of bias and study characteristics on accuracy of the four-symptom TB screen by including only studies that meet the following criteria in the meta-analysis.

- Studies that avoided case-control design.
- Studies that avoided inappropriate exclusions.
- Studies that interpreted the result of symptom screen (index test) without knowledge to the result of the reference standard.
- Studies where TB diagnosis is based solely on the reference of Xpert or culture (e.g. studies did not use CXR or clinical diagnosis as the basis for classifying patients as having “TB” or “No TB”).

Assessment of reporting bias

We will not formally assess reporting bias using funnel plots or regression tests as these have not been reported as helpful for diagnostic test accuracy studies (Macaskill 2010).

ACKNOWLEDGEMENTS

We thank Vittoria Lutje, Cochrane Infectious Diseases Group (CIDG) Information Specialist, and Hannah Rogers, Emory University, for their assistance with developing the search strategy.

The CIDG editorial base is supported by the Effective Health Care Research Consortium. This Consortium is funded by UK aid from the UK Government for the benefit of developing countries (Grant: 5242). The views expressed in this publication do not necessarily reflect UK government policy.

REFERENCES

Additional references

Ahmad Khan 2014

Ahmad Khan F, Verkuijl S, Parrish A, Chikwava F, Ntuny R, El-Sadr W, et al. Performance of symptom-based tuberculosis screening among people living with HIV: not as great as hoped. *AIDS* 2014;**28**(10):1463–72.

Bossuyt 2006

Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ* 2006;**332**(7549):1089–92.

Chu 2010

Chu H, Guo H, Zhou Y. Bivariate random effects meta-analysis of diagnostic studies using generalized linear mixed models. *Medical Decision Making* 2010;**30**(4):499–508.

Corbett 2003

Corbett EL, Watt CJ, Walker N, Maher D, Williams BG, Raviglione MC, et al. The growing burden of tuberculosis: global trends and interactions with the HIV epidemic. *Archives of Internal Medicine* 2003;**163**(9):1009–21.

Deluca 2009

Deluca A, Chaisson RE, Martinson NA. Intensified case finding for tuberculosis in prevention of mother-to-child transmission programs: a simple and potentially vital addition for maternal and child health. *Journal of Acquired Immune Deficiency Syndromes* 2009;**50**(2):196–9.

Getahun 2011

Getahun H, Kittikraisak W, Heilig CM, Corbett EL, Ayles H, Cain KP, et al. Development of a standardized screening rule for tuberculosis in people living with HIV in resource-constrained settings: individual participant data meta-analysis of observational studies. *PLoS Medicine* 2011;**8**(1): e1000391.

Getahun 2012

Getahun H, Sculier D, Sismanidis C, Grzemska M, Raviglione M. Prevention, diagnosis, and treatment of tuberculosis in children and mothers: evidence for action for maternal, neonatal, and child health services. *Journal of Infectious Diseases* 2012;**205**(Suppl 2):S216–27.

Gupta 2011

Gupta A, Chandrasekhar A, Gupte N, Patil S, Bhosale R, Sambarey P, et al. Symptom screening among HIV-infected

- pregnant women is acceptable and has high negative predictive value for active tuberculosis. *Clinical Infectious Diseases* 2011;**53**(10):1015–8.
- Hamadeh 1992**
Hamadeh MA, Glassroth J. Tuberculosis and pregnancy. *Chest* 1992;**101**(4):1114–20.
- Harris 2009**
Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* 2009;**42**(2):377–81.
- Hoffmann 2013**
Hoffmann CJ, Variava E, Rakgokong M, Masonoke K, van der Watt M, Chaisson RE, et al. High prevalence of pulmonary tuberculosis but low sensitivity of symptom screening among HIV-infected pregnant women in South Africa. *PLoS One* 2013;**8**(4):e62211.
- Horsburgh 2011**
Horsburgh CR, Rubin EJ. Clinical practice. Latent tuberculosis infection in the United States. *New England Journal of Medicine* 2011;**364**(15):1441–8.
- Kancheya 2014**
Kancheya N, Luhanga D, Harris JB, Morse J, Kapata N, Bweupe M, et al. Integrating active tuberculosis case finding in antenatal services in Zambia. *International Journal of Tuberculosis and Lung Disease* 2014;**18**(12):1466–72.
- Kourtis 2014**
Kourtis AP, Read JS, Jamieson DJ. Pregnancy and infection. *New England Journal of Medicine* 2014;**370**(23):2211–8.
- LaCourse 2016**
LaCourse SM, Cranmer LM, Matemo D, Kinuthia J, Richardson BA, John-Stewart G, et al. Tuberculosis case finding in HIV-infected pregnant women in Kenya reveals poor performance of symptom screening and rapid diagnostic tests. *Journal of Acquired Immune Deficiency Syndromes* 2016;**71**(2):219–27.
- Lawn 2011**
Lawn SD, Wood R, Wilkinson RJ. Changing concepts of “latent tuberculosis infection” in patients living with HIV infection. *Clinical and Developmental Immunology* 2011;**2011**:pii: 980594.
- Lawn 2012**
Lawn SD. Point-of-care detection of lipoarabinomannan (LAM) in urine for diagnosis of HIV-associated tuberculosis: a state of the art review. *BMC Infectious Diseases* 2012;**12**: 103.
- Lijmer 1999**
Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**(11):1061–6.
- Llewelyn 2000**
Llewelyn M, Cropley I, Wilkinson RJ, Davidson RN. Tuberculosis diagnosed during pregnancy: a prospective study from London. *Thorax* 2000;**55**(2):129–32.
- Macaskill 2010**
Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Chapter 10: Analysing and presenting results. *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy* Version 1.0.0. The Cochrane Collaboration, 2010. Available from srdata.cochrane.org. The Cochrane Collaboration.
- Mathad 2012**
Mathad JS, Gupta A. Tuberculosis in pregnant and postpartum women: epidemiology, management, and research gaps. *Clinical Infectious Diseases* 2012;**55**(11): 1532–49.
- Modi 2016**
Modi S, Cavanaugh JS, Shiraiishi RW, Alexander HL, McCarthy KD, Burmen B, et al. Performance of clinical screening algorithms for tuberculosis intensified case finding among people living with HIV in western Kenya. *PLoS One* 2016;**11**(12):e0167685.
- Modi 2017 [pers comm]**
Modi S (Centers for Disease Control and Prevention). [personal communication]. Conversation with: SM LaCourse (University of Washington) 21 March 2017.
- Nahid 2016**
Nahid P, Dorman SE, Alipanah N, Barry PM, Brozek JL, Cattamanchi A, et al. Official American Thoracic Society/ Centers for Disease Control and Prevention/Infectious Diseases Society of America Clinical Practice Guidelines: Treatment of drug-susceptible tuberculosis. *Clinical Infectious Diseases* 2016;**63**(7):e147–95.
- Pai 2016**
Pai M, Behr MA, Dowdy D, Dheda K, Divangahi M, Boehme CC, et al. Tuberculosis. *Nature Reviews Disease Primers* 2016;**2**:16076.
- Rangaka 2012**
Rangaka MX, Wilkinson RJ, Glynn JR, Boule A, van Cutsem G, Goliath R, et al. Effect of antiretroviral therapy on the diagnostic accuracy of symptom screening for intensified tuberculosis case finding in a South African HIV clinic. *Clinical Infectious Diseases* 2012;**55**(12):1698–706.
- Ratnapalan 2004**
Ratnapalan S, Bona N, Chandra K, Koren G. Physicians’ perceptions of teratogenic risk associated with radiography and CT during early pregnancy. *American Journal of Roentgenology* 2004;**182**(5):1107–9.
- Reitsma 2005**
Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**(10):982–90.

RevMan 2014 [Computer program]

Nordic Cochrane Centre, The Cochrane Collaboration. Review Manager 5 (RevMan 5). Version 5.3. Copenhagen: Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Saito 2012

Saito S, Howard AA, Reid M J, Elul B, Scardigli A, Verkuil S, et al. TB diagnostic capacity in sub-Saharan African HIV care settings. *Journal of Acquired Immune Deficiency Syndromes* 2012;**61**(2):216–20.

Salazar-Austin 2017

Salazar-Austin N, Hoffmann J, Cohn S, Mashabela F, Waja Z, Lala S, et al. Poor obstetric and infant outcomes in HIV-infected pregnant women with tuberculosis in South Africa: the Tshepiso study. *Clinical Infectious Diseases* 2017 Sep 26 [Epub ahead of print]. [DOI: 10.1093/cid/cix851]

Shah 2016

Shah M, Hanrahan C, Wang ZY, Dendukuri N, Lawn SD, Denkinger CM, et al. Lateral flow urine lipoarabinomannan assay for detecting active tuberculosis in HIV-positive adults. *Cochrane Database of Systematic Reviews* 2016, Issue 5. [DOI: 10.1002/14651858.CD011420.pub2]

Singh 2007

Singh N, Perfect JR. Immune reconstitution syndrome and exacerbation of infections after pregnancy. *Clinical Infectious Diseases* 2007;**45**(9):1192–9.

Somoskovi 2000

Somoskovi A, Kodmon C, Lantos A, Bartfai Z, Tamasi L, Fuzy J, et al. Comparison of recoveries of mycobacterium tuberculosis using the automated BACTEC MGIT 960 system, the BACTEC 460 TB system, and Lowenstein-Jensen medium. *Journal of Clinical Microbiology* 2000;**38**(6):2395–7.

StataCorp 2015 [Computer program]

StataCorp. Stata Statistical Software. Version Release 14. College Station, TX: StataCorp LP, 2015.

Steingart 2014

Steingart KR, Schiller I, Horne DJ, Pai M, Boehme CC, Dendukuri N. Xpert® MTB/RIF assay for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2014, Issue 1. [DOI: 10.1002/14651858.CD009593.pub3]

Stop TB Partnership 2015

Stop TB Partnership. The Paradigm Shift 2016-2020 Global Plan to End TB. www.stoptb.org/assets/documents/global/plan/GlobalPlanToEndTB_TheParadigmShift_2016-2020_StopTBPartnership.pdf (accessed 1 February 2017).

Sugarman 2014

Sugarman J, Colvin C, Moran A, Oxlade O. Tuberculosis in pregnancy: an estimate of the global burden of disease. *Lancet* 2014;**2**(12):e710–6.

Takwoingi 2017

Takwoingi Y, Guo B, Riley RD, Deeks JJ. Performance of methods for meta-analysis of diagnostic test accuracy with few studies or sparse data. *Statistical Methods in*

Medical Research 2017;**26**(4):1896–911. [DOI: 10.1177/0962280215592269]

UNAIDS 2016

UNAIDS. Global HIV Statistics Fact Sheet. www.unaids.org/en/resources/fact-sheet. UNAIDS, (accessed 1 February 2017).

UNITAID 2015

UNITAID. Tuberculosis Diagnostics Technology and Market Landscape 4th edition. unitaid.org/images/marketedynamics/publications/Tuberculosis_diagnostics_technology_and_market_landscape_4th_edition_Oct_2015.pdf (accessed 1 March 2017).

Whiting 2011

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529–36.

WHO 2010

World Health Organization. Treatment of tuberculosis guidelines. Fourth Edition. apps.who.int/iris/bitstream/10665/44165/1/9789241547833_eng.pdf?ua=1&ua=1. World Health Organization, (accessed 23 March 2017).

WHO 2011a

World Health Organization. Guidelines for intensified tuberculosis case-finding and isoniazid preventive therapy for people living with HIV in resource-constrained settings. whqlibdoc.who.int/publications/2011/9789241500708_eng.pdf. World Health Organization, (accessed 1 February 2017).

WHO 2011b

World Health Organization. Policy statement: Automated real-time nucleic acid amplification technology for rapid and simultaneous detection of tuberculosis and rifampicin resistance: Xpert MTB/RIF system. whqlibdoc.who.int/publications/2011/9789241501545_eng.pdf. World Health Organization, (accessed 1 February 2017).

WHO 2013

World Health Organization. Systematic screening for active tuberculosis: principles and recommendations. www.who.int/tb/tbscreening/en/. World Health Organization, (accessed 1 February 2017).

WHO 2015a

World Health Organization. Guideline on when to start antiretroviral therapy and on pre-exposure prophylaxis for HIV. apps.who.int/iris/bitstream/10665/186275/1/9789241509565_eng.pdf (accessed 26 April 2017).

WHO 2015b

World Health Organization. The use of lateral flow urine lipoarabinomannan assay (LF-LAM) for the diagnosis and screening of diagnosis and screening of active tuberculosis in people living with HIV. Policy update. www.who.int/tb/publications/use-of-lf-lam-tb-hiv/en/. World Health Organization, (accessed 21 March 2017).

WHO 2015c

World Health Organization. Recommendation on 36 months isoniazid preventive therapy to adults and adolescents living with HIV in resource-constrained and high TB and HIV-prevalence settings: Update. apps.who.int/iris/bitstream/10665/174052/1/9789241508872_eng.pdf?ua=1&ua=1. World Health Organization, (accessed 24 March 2017).

World Bank 2016

World Bank. World Bank list of economies.

datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups. World Bank, (accessed 4 March 2017).

Zenner 2012

Zenner D, Kruijshaar ME, Andrews N, Abubakar I. Risk of tuberculosis in pregnancy: a national, primary care-based cohort and self-controlled case series study. *American Journal of Respiratory and Critical Care Medicine* 2012;**185**(7):779–84.

* Indicates the major publication for the study

APPENDICES

Appendix I. Search strategy

We will include search terms to capture clinical tuberculosis symptoms among HIV-infected pregnant women. We will utilize both MeSH terms and free text terms to identify phrases of our search. In order to increase the sensitivity of our findings, we will search “All fields” when possible rather than only title and abstracts. We will include the following search terms.

1 exp HIV Infections/

2 HIV/ or HIV-2/ or HIV-1/

3 (HIV or hiv-1* or hiv-2* or hiv1 or hiv2 or hiv-I* or hiv-II* or hivI*).tiab

4 “HIV infect*” .tiab

5 (“human immunodeficiency virus” or “human immunodeficiency virus” or “human immuno-deficiency virus” or “human immunodeficiency virus” or (human immun* and deficiency virus)).tiab

6 (“acquired immunodeficiency syndrome” or “acquired immunodeficiency syndrome” or AIDS).tiab

7 (“acquired immuno-deficiency syndrome” or “acquired immune-deficiency syndrome”).tiab

8 or/1-7

9 exp Tuberculosis/

10 (tuberculosis or TB or tuberculoses or tuberculous).tiab

11 Mycobacterium tuberculosis/

12 9 or 10 or 11

13 8 and 12

14 (pregnant adj3 wom?n).mp.

15 Pregnant Women/

16 Pregnancy/

17 pregnan*.tiab

18 14 or 15 or 16 or 17

19 13 and 18

20 mass screening.mp. or Mass Screening/

21 physical examination.mp. or Physical Examination/

22 Cough/

23 Weight Loss/

24 Fever/ or Fatigue/

25 Prenatal Diagnosis/ or ((prenatal* or pre natal* or antenatal* or ante natal*) adj2 screen*).tiab.

26 clinical algorith*.mp.

27 (cough* or “weight loss*” or “weight reduction*” or fever* or pyrexia* or “night sweat*” or fatigue*).tiab

28 (“four-symptom screen*” or “four-symptoms screen*”).tiab

29 triag*.mp. or Triage/

30 case finding.mp.

31 symptom* adj2 screen* .mp.
 32 or/20-31
 33 19 and 32

This is the preliminary search strategy for MEDLINE. It will be adapted for other electronic databases. All search strategies will be reported in full in the final version of the review.

Appendix 2. Data extraction form

Symptom screen for active tuberculosis in pregnant women living with HIV: data extraction form	
Study ID	
Name of data extractor	1 - SL 2 - JM 3 - Other (Specify _____)
First Author	
Corresponding author and email	
Was an author contacted?	1 - Yes Contact Date: _____ <input type="checkbox"/> YES Response Result: <input type="checkbox"/> NO Response 2 - No
Title of study	
Year of publication	
Year(s) study conducted	
Language	1 - English 2 - French 3 - Spanish 4 - Other: _____
Journal	
Publication status of study	1 - Published 2 - Unpublished What is the anticipated study completion date?
Country or countries where study was conducted (list all)	
World Bank Classification	1 - Middle/Low 2 - High 3 - Study includes both middle/low & high
For the diagnosis of active TB, what reference standard was used to identify TB and not TB? Circle all that apply NAAT - nucleic acid amplification test LJ - Lowenstein-Jensen ZN - Ziehl-Neelsen stain	1 - Sputum: Solid Culture (circle method) LJ 7H10 7H11 2 - Sputum: Liquid Culture (circle method) MGIT Bactec460 3 - Sputum: Both Solid and Liquid Culture (specify above) 4 - Sputum: NAAT (circle method) GeneXpert Other (specify): _____ 5 - Sputum: Smear (circle method) ZN FM 6 - Sputum: Other, specify: _____

(Continued)

FM - Fluorescence microscopy	7 - Non-sputum: specify: _____ 9 - Unknown/not reported
Were patients unable to produce sputa included in this study?	1 - Yes 2 - No 9 - Unknown/not reported
How many sputum specimens per patient were obtained for the diagnosis of active TB?	1 - One 2 - Two 3 - Other (specify): _____ 9 - Unknown/not reported
What was the clinical setting of the study?	1 - Outpatient, Maternal/child health clinic 2 - Outpatient, HIV clinic or other general medicine clinic 3 - Inpatient 4 - Both out-patient and in-patient 5 - Community-based 6 - Other, describe: _____ 9 - Unknown/not reported
Study design	1 - Randomized trial 2 - Cross-sectional 3 - Cohort 4 - Case-control 5 - Other (specify): _____ 9 - Unknown/not reported
What was the manner of patient selection into the study?	1 - Consecutive 2 - Random 3 - Convenience 4 - Other, specify 9 - Unclear/not reported
Direction of study data collection	1 - Prospective 2 - Retrospective 9 - Unknown/not reported
Comments about study design	
Index tests: What symptoms were evaluated as screens for TB? Indicate all that apply	1 - Cough 2 - Fever 3 - Night sweats 4 - Weight loss 5 - Cough > 2 week duration 6 - Failure to gain weight 7 - Other, specify: _____

(Continued)

Please select the statement that best describes the purpose of screening as described in the study	1- Pregnant PLHIV with signs or symptoms suggestive of active TB were screened for TB only 2 - Pregnant PLHIV with and without TB signs or symptoms were screened for TB 3 - Neither 1 nor 2. 4 - Other, specify: _____	
Number after screening by exclusion & inclusion criteria (Total for study)	____ 9 - Unknown/not reported	
Number included in analysis (# screened - # withdrawals) (Total for study)	____ 9 - Unknown/not reported	
HIV status of participants (Total for study)	1 - HIV-positive 2 - Both HIV-positive and HIV-negative 9 - Unknown/not reported	
Percentage of HIV-negative and HIV-positive	____ % of participants were HIV-negative ____ % of participants were HIV-positive Specify numerator/denominator	
Pregnancy status of participants	1 - Pregnant 2 - Postpartum 3 - Pregnant and postpartum 9 - Unknown/not reported	
Percentage of pregnant and postpartum	____ % of participants were pregnant ____ % of participants were postpartum Specify numerator/denominator	
Gestational or Postpartum age	Pregnant: Range:_____ Mean (SD):_____ ____ Median (IQR):_____ ____ 9 - Unknown/not reported	Postpartum: Range:_____ Mean (SD):_____ Median (IQR):_____ 9 - Unknown/not reported
Percentage of pregnant and/or postpartum and HIV-positive	____ % of participants were pregnant or postpartum, and HIV-positive Specify numerator/denominator	
Age of participants	All Study Participants: Range:_____ Mean (SD):_____ ____ Median (IQR):_____ ____	Pregnant PLHIV: Range:_____ Mean (SD):_____ Median (IQR):_____ 9 - Unknown/not reported

(Continued)

	<p>—</p> <p>9 - Unknown/not reported</p>	
Reported AFB-smear status	<p>All Study Participants:</p> <p>— % of participants have AFB-smear positive pulmonary TB</p> <p>Specify numerator/denominator.——</p> <p>9 - Unknown/not reported</p>	<p>Pregnant PLHIV:</p> <p>— % of participants have AFB-smear positive pulmonary TB</p> <p>Specify numerator/denominator.——</p> <p>9 - Unknown/not reported</p>
Participant engagement in medical care (indicate all that apply)	<p>All Study Participants:</p> <p>1 - Previously engaged in HIV care</p> <p>2 - Never before received HIV care</p> <p>3 - Both</p> <p>4 - Other, specify: —</p> <p>————</p> <p>9 - Unknown/not reported</p>	<p>Pregnant PLHIV:</p> <p>1 - Previously engaged in HIV care</p> <p>2 - Never before received HIV care</p> <p>3 - Both</p> <p>4 - Other, specify: —</p> <p>9 - Unknown/not reported</p>
Proportion of participants with HIV never before engaged in care (i.e. enrolling into HIV care)	<p>PLHIV:</p> <p>— % of participants previously engaged in care</p> <p>— % of participants never before engaged in care</p> <p>Specify numerator/denominator.——</p> <p>9 - Unknown/not reported</p>	<p>Pregnant PLHIV:</p> <p>— % of participants previously engaged in care</p> <p>— % of participants never before engaged in care</p> <p>Specify numerator/denominator.——</p> <p>9 - Unknown/not reported</p>
ART status of PLHIV	<p>PLHIV:</p> <p>1 - No ART</p> <p>2 - Combined ART</p> <p>3 - ART for PMTCT only</p> <p>9 - Unknown/not reported</p>	<p>Pregnant PLHIV:</p> <p>1 - No ART</p> <p>2 - Combined ART</p> <p>3 - ART for PMTCT only</p> <p>9 - Unknown/not reported</p>
Proportion of PLHIV participants receiving combination ART (i.e. not only treated for PMTCT)	<p>——% of participants were on no ART</p> <p>——% of participants were on combined ART</p> <p>——% of participants were on ART for PMTCT only</p> <p>9 - Unknown/not reported</p>	

(Continued)

What was the CD4-cell count of included patients?	PLHIV: Range:_____ Mean (SD):_____ _____ Median (IQR):_____ _____ 9 - Unknown/not reported	Pregnant PLHIV: Range:_____ Mean (SD):_____ Median (IQR):_____ 9 - Unknown/not reported		
What was the HIV viral load of included patients?	PLHIV: _____ % with undetectable viral load Range:_____ Mean (SD):_____ _____ Median (IQR):_____ _____ 9 - Unknown/not reported	Pregnant PLHIV: _____ % with undetectable viral load Range:_____ Mean (SD):_____ Median (IQR):_____ 9 - Unknown/not reported		
Did the study include patients with previous TB history?	All Study Participants: 1 - Yes _____ % of participants had a prior history of TB Specify numerator/denominator_____ 2 - No 9 - Unknown/not reported	Pregnant PLHIV: 1 - Yes _____ % of participants had a prior history of TB Specify numerator/denominator_____ 2 - No 9 - Unknown/not reported		
Reference Test: Xpert or culture				
WHO four-symptom screen		TB Disease	Not TB	Total
	Positive			
	Negative			
	Total			
	Reference Test: Xpert or culture			
	Pregnant PLHIV on ART			

(Continued)

WHO screen	four-symptom		TB Disease	Not TB	Total
		Positive			
		Negative			
		Total			
		Pregnant PLHIV not on ART	Reference Test: Xpert or culture		
WHO screen	four-symptom		TB Disease	Not TB	Total
		Positive			
		Negative			
		Total			
Symptom (specify symptom):			Reference Test: Xpert or Culture		
			TB Disease	Not TB	Total
			Positive		
			Negative		
			Total		

Appendix 3. QUADAS-2

Domain 1: patient selection

Risk of bias: could the selection of patients have introduced bias?

Signalling question 1: Was a consecutive or random sample of patients enrolled?

We will score 'yes' if the study enrolled a consecutive or random sample of eligible participants (i.e., pregnant PLHIV); 'no' if the study selected participants by convenience; and 'unclear' if the study did not report the manner of participant selection or we could not tell.

Signalling question 2: Was a case-control design avoided?

Case control study design may overestimate sensitivity and specificity for screening and diagnostic tests (Lijmer 1999). We will score 'yes' to studies which are not case-control studies. We will score 'no' to studies which are case-control studies. We will score 'unclear' if we could not tell.

Signalling question 3: Did the study avoid inappropriate exclusions?

We will score 'yes' to studies, which included: a) all pregnant PLHIV regardless of symptoms and b) pregnant PLHIV who were unable to produce sputum (expectorated or induced). We will score 'no' if studies excluded pregnant PLHIV on the basis of no symptoms or the inability to produce sputum (no attempts at sputum induction). We will also score 'no' if studies excluded pregnant PLHIV presumed to have extrapulmonary TB. We will score 'unclear' if we could not tell.

Applicability: Are there concerns that the included patients and setting do not match the review question?

We are interested in how the four-symptom TB screen performs in pregnant PLHIV who would be screened in routine practice. We have defined 'screening' for active TB in accordance with WHO guidance, as "the systematic identification of people with suspected active TB, in a predetermined target group, using tests, examinations or other procedures that can be applied rapidly" (WHO 2013). We will score 'low concern' for studies in which the four-symptom TB screen was performed uniformly within the predetermined study target population of pregnant PLHIV, 'high concern' if the four-symptom TB screen was not performed uniformly within the predetermined study target population of pregnant PLHIV, and 'unclear concern' if we could not tell.

Domain 2: index test**Risk of bias: could the conduct or interpretation of the index test have introduced bias?****Signalling question 1: were the index test results interpreted without knowledge of the results of the reference standard?**

We will score 'yes' if the study interpreted the result of the four-symptom TB screen blinded to the result of the reference standard or if it is clear that the results of the index test were available before the results of the reference standard were known; we will score 'no' if the study did not interpret the result of the four-symptom TB screen blinded to the result of the reference standard. We will score 'unclear' if we could not tell if the index test results were interpreted without knowledge of the reference standard results. We will also answer yes if the tests (index test and reference standard) were carried out in different places.

Signalling question 2: if a threshold was used to define positivity, was it prespecified?

This question is not applicable for our review.

Applicability: are there concerns that the index test, its conduct, or its interpretation differ from the review question?

If index test methods vary from those specified in the review question, concerns about applicability may exist. We will score 'high concern' if the four-symptom TB screen was applied for the purpose of TB diagnosis, rather than as a TB screening tool; 'low concern' if the four-symptom TB screen was applied as a screening tool, and 'unclear concern' if we could not tell.

Domain 3: reference standard**Risk of bias: could the reference standard, its conduct, or its interpretation have introduced bias?****Signalling question 1: is the reference standard likely to correctly classify the target condition?**

Microbiological reference standard

Mycobacterial culture (liquid or solid) or nucleic acid amplification tests (e.g., Xpert) are considered the best reference standards to identify active TB in PLHIV. Due to the difficulties in diagnosing HIV-associated TB, it is recommended that multiple cultures from sputum be evaluated.

We will answer 'no' if a consistent approach was not followed for all patients (for example, some but not all patients were asked to provide sputum for Xpert or culture testing). We will answer 'unclear' if we could not tell.

Signalling question 2: were the reference standard results interpreted without knowledge of the results of the index test?

We will answer 'yes' if the study interpreted the result of the reference standard blinded to the result of the four-symptom TB screen, or if the reference standard result was reported on an automated instrument; 'no' if the study did not interpret the result of the reference standard blinded to the result of the four-symptom TB screen, and 'unclear' if we could not tell. We will also answer yes if the tests carried out in different places.

Applicability: are there concerns that the target condition as defined by the reference standard does not match the question?

In general, we think there will be low concern for almost included studies based on the current definition of the reference standard. We will judge 'high concern' if the included studies did not speciate mycobacteria isolated in culture, 'low concern' if speciation was performed, and 'unclear' if we could not tell.

Domain 4: Flow and timing

Risk of bias: could the patient flow have introduced bias?

Signalling question 1: was there an appropriate interval between the index test and reference standard?

We will answer 'yes' if the index test and reference standard(s) are collected on the same patients at the same time or within seven days. We chose seven days as a time period during which either treatment of TB or natural progression of TB without treatment could impact test results. We will answer 'no' if specimens were collected for index and reference standard tests greater than seven days apart, and 'unclear' if we could not tell.

Signalling question 2: did all patients receive the same reference standard?

We will answer 'yes' if all participants in the study received the same reference standard to confirm TB; 'no' if not all patients received the reference standard to confirm TB, and 'unclear' if we could not tell.

Signalling question 3: were all patients included in the analysis?

We will determine the answer to this question by comparing the number of participants enrolled in the study with the number of participants included in the two-by-two tables. We will answer 'yes' if the number of participants in the two-by-two tables match the number of participants recruited into the study, or if these numbers do not match, then sufficient explanation is provided for any discrepancy. We will answer 'no' if the number of participants in the two-by-two tables do not match the number of participants recruited into the study and insufficient explanation is provided for any discrepancy, and 'unclear' if we could not tell.

Judgments for overall 'Risk of bias' assessments for domains

If we answer:

- all signalling questions for a domain "yes," then we will judge risk of bias "low";
- all or most signalling questions for a domain "no," then we will judge risk of bias "high";
- one signalling question for a domain "no," we will discuss with a third author the 'Risk of bias' judgement;
- all or most signalling questions for a domain "unclear," then we will judge risk of bias "unclear";
- only one signalling question for a domain "unclear," we will discuss with a third author the 'Risk of bias' judgement for the domain.

CONTRIBUTIONS OF AUTHORS

SL wrote the initial draft of the protocol. LC developed the search strategy. DH designed draft abstraction forms. EO drafted the data analysis sections with input from SP and KRS. KRS drafted the QUADAS-2 section. KRS and SP contributed methodological advice. All authors (EO, AB, DB, LC, DH, JM, SL, SM, KRS, and SP) provided input for the protocol.

DECLARATIONS OF INTEREST

The review authors (EO, AB, DB, LC, DH, JM, SL, SM, KRS, SP) have no known conflicts of interest. LC, DH, SL, JM, and SM have conducted primary research regarding evaluating symptom screening for TB in pregnant PLHIV.

SOURCES OF SUPPORT

Internal sources

- Liverpool School of Tropical Medicine, UK.

External sources

- Cystic Fibrosis Foundation (DH), USA.
- NIH/NIAID K23 Career Development Award (SL, DH), USA.
- NIH/NIAID R21 (DH), USA.
- University of Washington Center for AIDS Research (SL, DH), USA.
- University of Washington Strategic Analysis, Research & Training (START) Center (DB), USA.
- US President's Emergency Plan for AIDS Relief through the US Centers for Disease Control and Prevention (SM, SP), USA.
- NIH/NIAID K23 (K23AI129854) Award (JM), USA.
- NIH/NICHD R01 (R01HD081929) Award (JM), USA.
- Department for International Development (DFID), UK.

Grant: 5242