INT J TUBERC LUNG DIS 23(6):728–734

http://dx.doi.org/10.5588/ijtld.18.0547

Screening tools to exclude active pulmonary TB in high TB burden countries: systematic review and meta-analysis

Y. Assefa,* S. Woldeyohannes,* Y. A. Gelaw,* Y. Hamada,† H. Getahun†

***School of Public Health, the University of Queensland, Brisbane, Queensland, Australia; †Global Tuberculosis Programme, World Health Organization, Geneva, Switzerland**

SUMMARY

OBJECTIVE: To examine the use of symptoms, chest X-ray (CXR) abnormalities, and combinations of symptoms and CXR in excluding active pulmonary tuberculosis (TB) before treating for latent tuberculous infection (LTBI) in high TB burden countries.

METHODS: We updated a systematic review and meta-analysis of studies on the sensitivities, specificities, predictive values, diagnostic odds ratios and areas under the curve for index tests. The analysis was conducted using the hierarchical summary receiver operating characteristic method in R software.

RESULTS: We included 24 publications in the systematic review and meta-analysis. 'Any CXR abnormality' had the highest sensitivity (94.1%, 95%CI 85.8–97.7) among all index tests. 'CXR abnormality suggestive of TB' had a higher specificity (92.2%, 95%CI 89.7–94.1) than 'any CXR abnormality' (86.8%, 95%CI 79.7–91.7). The sensitivity for 'any TB symptom' was 73.0% (95%CI 64.1–80.4), while 'prolonged cough' of ≥2 weeks had a specificity of 94.3% (95%CI 92.2–95.9). There was no significant difference in the sensitivity and specificity of all screening tools stratified by human immunodeficiency virus (HIV) settings, with the exception of 'CXR abnormality suggestive of TB', which had a significantly higher sensitivity in low than in high HIV prevalence settings (effect estimate 2.26, 95%CI 0.69–3.82; $P = 0.002$).

CONCLUSION: In countries with a high TB burden, the absence of any TB symptom and any CXR abnormality can be used to exclude active pulmonary TB before initiating treatment for LTBI in household contacts aged ≥5 years of patients with bacteriologically confirmed pulmonary TB.

KEY WORDS: tuberculosis; screening tools, excluding active tuberculosis; latent tuberculous infection; high tuberculosis countries

ONE THIRD TO ONE FOURTH of the world's population is estimated to be infected with *Mycobacterium tuberculosis*.[1,2] However, a relatively small proportion (5–10%) go on to develop TB disease during their lifetime. The remaining 90–95% of infected people have latent tuberculous infection (LTBI), which may be reactivated at a later stage, with a higher risk in the presence of conditions that affect immunity.[3] Management of LTBI is one of the strategies stated in the World Health Organization's (WHO's) End TB Strategy.[4] LTBI management involves testing and delivering effective treatment whose benefits are considered to outweigh the harms.

The 2015 WHO guidelines on the programmatic management of LTBI are intended to help countries accelerate the implementation of measures and expand activities for LTBI management. The guidelines were primarily targeted at low TB burden countries, which are arbitrarily defined as high- or upper middle-income countries with an estimated TB incidence rate of <100 per 100 000 population. Due to their high sensitivity, the guidelines recommended the use of a combination of chest radiography (CXR) with any abnormality or any TB symptom to exclude active TB before commencing LTBI treatment.[5] These recommendations were based on a systematic review conducted in 2012 to assess the sensitivity and specificity of screening tools (symptoms, CXR abnormalities, and combinations of symptoms and CXR) for detecting bacteriologically confirmed active pulmonary TB (PTB) in human immunodeficiency virus (HIV) negative persons and persons with unknown HIV status considered eligible for TB screening.[6]

However, preventive treatment of LTBI in household contacts aged ≥5 years in high TB burden countries had not been recommended. Moreover, there were no recommendations for an algorithm to exclude active TB before considering LTBI treatment among household contacts aged ≥5 years in high TB burden countries. After repeated requests from several countries and consultations with stakehold-

Correspondence to: Yibeltal Assefa, School of Public Health, University of Queensland, 266 Herston Rd, Herston, QLD 4006, Australia. e-mail: yibeltalassefa343@gmail.com; y.alemu@uq.edu.au

Article submitted 8 August 2018. Final version accepted 27 November 2018.

ers, the WHO recognised the need for new guidance on LTBI management, including the importance of LTBI testing and treatment among HIV-negative children aged >5 years, adolescents and adults who are contacts of patients with PTB.

The present review was conducted to inform the WHO in defining an algorithm for excluding active PTB among household contacts aged ≥5 years of patients with PTB in high TB burden countries.

We aimed to develop an algorithm that would help countries exclude active TB before initiating LTBI treatment. We conducted an update of the earlier systematic review on screening tools for excluding active PTB; our review was restricted to studies from high TB burden countries. The current review informed the development of the recently updated and consolidated guidelines for the programmatic management of LTBI defined by the WHO.[7]

METHODS

Eligibility criteria

This review included only cross-sectional studies. We excluded studies that only identified persons with tuberculous infection but not active TB. Our review was limited to studies in English, French and Spanish. Studies were selected according to the criteria outlined below based on population, index test, reference test and diagnosis of interest.[8]

Population

The primary focus of our review was people aged ≥5 years in contact with PTB cases. However, the review also included studies in individuals without TB contacts, such as the general population, as contacts are considered otherwise healthy individuals. We also included studies that combined adults and children.

Index test(s)

The index test of interest, which was evaluated for accuracy, was symptom screen or CXR, or symptom screen plus CXR combined sequentially or in parallel.[9] In sequential screening, individuals were screened for symptoms first, followed by CXR screening of only those who presented with symptoms. Parallel screening included both symptom and CXR screening conducted simultaneously.[10]

Symptom screening included any or combinations of the following symptoms suggestive of TB: cough, weight loss, haemoptysis, fatigue, fever/chills, night sweats, difficulty breathing/shortness of breath, pain/tightness in chest, loss of appetite and lymphadenopathy. Prolonged cough (also called 'chronic cough') was defined as persistent cough of ≥2 weeks.[11]

CXR involved posterior-anterior CXR recorded using conventional CXR and digital radiography. Studies using computer-assisted CXR were excluded from this review. We classified CXR findings as 'no abnormality', 'any abnormality' or 'abnormality suggestive of TB' as defined by study authors.[12,13]

Reference test(s)

Reference tests included mycobacterial culture (solid or liquid medium), and/or sputum smear microscopy, and/or Xpert® MTB/RIF (Cepheid, Sunnyvale, CA, USA).[14]

Diagnosis of interest

The diagnosis of interest was confirmed active PTB characterised by the presence of *M. tuberculosis* in sputum confirmed using the reference test defined above.

Search strategy and selection of publications

We conducted a literature search using criteria determined a priori in the MEDLINE, EMBASE (Excerpta Medica dataBASE), LILACS (Latin American & Caribbean Health Sciences Literature) and HTA (Health Technology Assessment) online databases to identify peer-reviewed articles published from August 2012 to January 2017. The search terms were combinations of the following three domains: 1) 'tuberculosis' and related terms, 2) 'screening', 'survey', 'sensitivity', 'specificity', and related terms, and 3) 'bacterial', 'culture', and 'microscopy'.

We used a three-step selection process. During the first stage, one reviewer (SW) excluded articles based on their titles. In the second stage, two reviewers (YA and SW) reviewed the abstracts of all retained studies independently and in duplicate. Finally, two reviewers (YA and SW) assessed the full text of every article that was potentially relevant independently and in duplicate.

Collection, extraction and management of data

Two reviewers (YA and SW) independently and in duplicate extracted data from all included sources using a pre-prepared and piloted data extraction form. Extracted information included study methodology; study setting; study population and participant demographics and baseline characteristics; details of the index and references tests; and outcome measurements such as sensitivity, specificity and predictive values. The standards for the reporting of diagnostic accuracy studies (STARD) checklist and a flow diagram was used as a detailed guide for data extraction.[15]

Risk of bias in individual studies

We used Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) to assess the quality of diagnostic accuracy studies.[16] Critical appraisal questions in QUADAS-2 were answered as 'Yes', 'No', 'Unclear' or 'Not applicable'. The appraisal was independently conducted by two reviewers (YA and SW); disagreements were resolved through discussion or by seeking the opinion of a third reviewer. The

results of the critical appraisal were presented in a table showing the results of the critical appraisal, and in a narrative summary of the overall methodological quality of the studies included.[17] If a study was judged as 'low' on all domains relating to bias or applicability, an overall judgment of 'low risk of bias' or 'low concern regarding applicability' was assigned to the study. If a study was judged 'high' or 'unclear' in one or more domains, it was judged as 'at risk of bias' or as having 'concerns regarding applicability'.[16]

Data analysis

Diagnostic two-by-two tables were generated, from which sensitivities and specificities for each index test with 95% confidence intervals (CIs) were calculated. Summary tables were prepared on the following: number screened, prevalence, TB cases, sensitivity, specificity, number of true-negatives (TN), false-negatives (FN), true-positives (TP), false-positives (FP), negative predictive value (NPV; i.e., the probability that subjects with a negative screening test definitely do not have the disease after screening and screening + confirmatory test), positive predictive value (PPV; i.e., the probability that subjects with a positive screening test truly have the disease after positive screening). Special emphasis was placed on NPV, as the objective of the review was to determine an algorithm to identify people without active PTB. A summary table was also produced using 1% and 5% TB prevalence.

We calculated the pooled sensitivity and specificity. However, as the use of paired indicators can be a disadvantage when comparing test performance, we also calculated diagnostic odds ratios (DORs; defined as the ratio of the odds of the test being positive if the subject has a disease relative to the odds of the test being positive if the subject does not have the disease) as a single indicator of diagnostic performance. We also constructed the respective summary receiver operating characteristic curves (SROCs) using the true-positive rate (sensitivity) plotted as a function of the false-positive rate (1–specificity) for different cut-off points of a parameter, and calculated the area under the curve (AUC). Although the SROC model is simple and adequate for summarising paired estimates of sensitivity and (1-specificity) across studies, it is unable to distinguish within-study and between-study variability, which results in biased estimates. The hierarchical summary receiver operating characteristic (HSROC) method was thus used to estimate overall diagnostic accuracies taking into account covariates.[9] Heterogeneity was assessed by visual inspection of forest plots and SROC. We also tested heterogeneity using the Cochrane Q-test (for DOR) and by determining the I^2 metric.[18] As HIV prevalence in countries can cause heterogeneity in sensitivity and specificity, we performed meta-regression analysis using a bivariate model including HIV prevalence as a covariate at the study level. Statistical significance was defined as $P < 0.05$. We used Endnote X5 (Clarivate Analytics, Philadelphia, PA, USA) to manage references, and Review Manager v5.3 (Cochrane Community, London, UK) and R Software v3.4.2 (R Computing, Vienna, Austria) to conduct meta-analysis.[19,20]

RESULTS

The current review identified 3201 unique hits. We excluded 1239 duplicates and 1848 articles by title and abstract review. An additional 108 articles were excluded because of an inappropriate population, index test or reference test. Finally, six articles were found to be eligible for updating the systematic review (Figure 1). The search criteria are also presented in Supplementary Data Section 1.

The previous review included 24 publications from 17 studies. We excluded six studies and publications because they were conducted on special population groups (five studies and publications) or were from low TB burden countries (one publication). We thus included 18 publications from 11 studies from the previous review in the updated systematic review. As some of the studies resulted in more than one publication, the number of publications ($n = 24$) exceeded the number of studies ($n = 17$) (Figure 1 and Supplementary Table S1).

General characteristics of included studies

The general characteristics of the studies included in this review are given in Supplementary Table S1. All studies except one involved people aged ≥15 years. We arbitrarily classified study settings as 'high' and 'low' HIV prevalence if the HIV prevalence in the general population was ≥5%. Twelve publications (eight studies) were from high HIV prevalence settings, while the remaining 12 (nine studies) were from low HIV prevalence settings. The risk of bias and concerns with applicability in individual studies ranged from low ($n = 7$) to high ($n = 7$). A few studies also had an unclear risk of bias and concerns with applicability ($n = 3$). The results of the quality assessment are given in the Supplementary Data Section 2.

Pooled estimates from index tests

Supplementary Table S2 shows that screening for TB using CXR resulted in a significantly higher sensitivity than screening with any of the symptom screening tools.

Any CXR abnormality

Seven studies evaluated the accuracy of 'any CXR abnormality' (Supplementary Table S2). The sensitivity ranged from 72% (95% confidence interval [CI] 64–79) to 99% (95%CI 98–100), and the

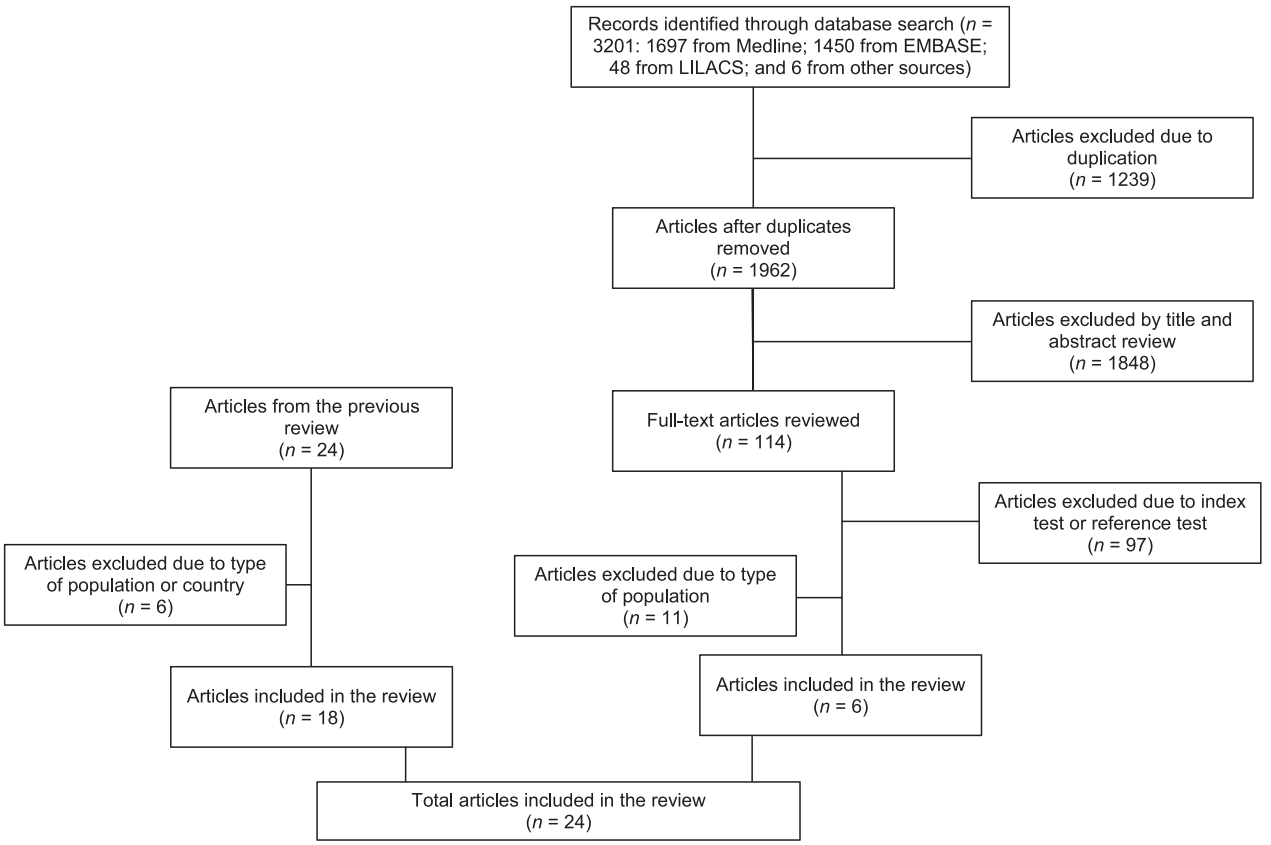

Figure 1 Flow diagram of literature search and selection strategy. EMBASE = Excerpta Medica dataBASE; LILACS = Latin American & Caribbean Health Sciences Literature.

specificity ranged from 73% (95%CI 72–73) to 94% (95%CI 94–95) (Figure 2). The pooled DOR was 45.9 (95%CI 10.2–206.5). There was no significant difference in DOR of the included studies with Cochran's Q and Higgins's I^2 (3.8 vs. 0%; $P = 0.427$).

The bivariate diagnostic random-effects meta-analysis indicated that the AUC was 0.95. There was no significant difference in sensitivity of any CXR abnormality in low and high HIV prevalence settings (coefficient 1.17, 95%CI –0.15 to 2.48; $P = 0.077$). There was also no significant difference in the specificity of any CXR abnormality in low and high HIV prevalence settings (coefficient 0.67, 95%CI –1.05 to 2.39; $P = 0.445$).

CXR with abnormalities suggestive of tuberculosis

Of the seven studies reporting 'CXR with abnormalities suggestive of TB', six used bacteriologically confirmed TB as the reference standard (Supplementary Table S2). The sensitivity ranged from 74% (95%CI 65–82) to 97% (95%CI 94–99), and the specificity ranged from 87% (95%CI 85–89) to 96% (95%CI 96–96) (Supplementary Figure S1). The pooled estimates of sensitivity and specificity were

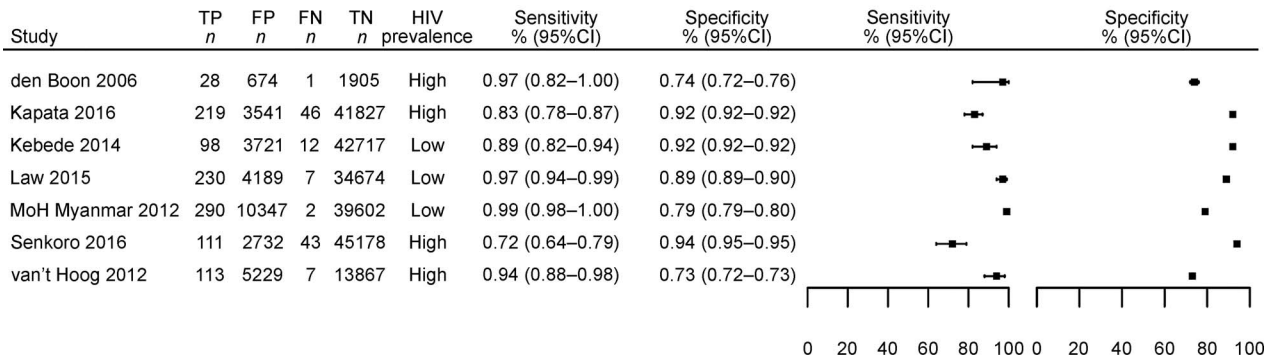

Figure 2 Forest plot of the sensitivity and specificity of CXR screening for any abnormality in the general population. TP = true-positive; FP = false-positive; FN = false-negative; TN = true-negative; HIV = human immunodeficiency virus; CI = confidence interval; MOH = Ministry of Health; CXR = chest X-ray.

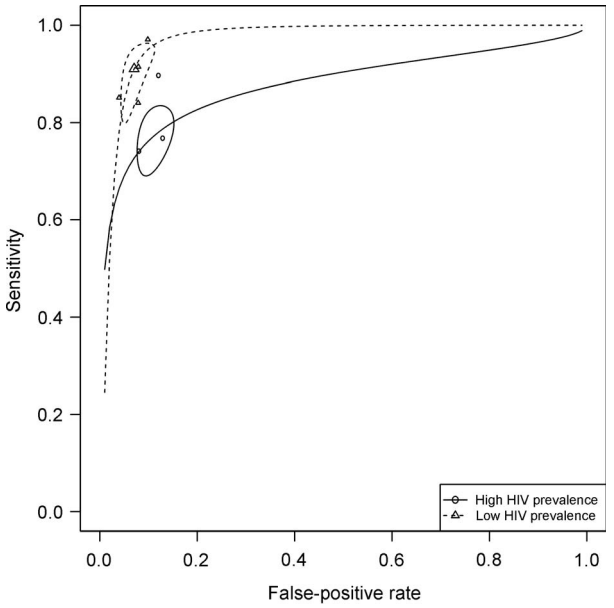

Figure 3 Summary receiver operating characteristic plots using CXR with abnormalities suggestive of TB as a screening tool in high and low HIV prevalence settings. HIV = human immunodeficiency virus; CXR = chest X-ray; TB = tuberculosis.

respectively 89.3% (95%CI 81.4–94.1) and 92.2% (95%CI 89.7–94.1).

The pooled DOR was 25.5 (95%CI 10.1–64.8). There was no significant difference in DOR of the included studies with Cochran's *Q* and Higgins's I^2 (14.3 vs. 15.9%; $P = 0.284$). The bivariate diagnostic random-effects meta-analysis indicated that the AUC was 0.9. Figure 4 gives the ROC curves by HIV prevalence. There was a significant difference in the sensitivity of CXR with abnormalities suggestive of TB in low compared with high HIV prevalence settings (coefficient 2.29, 95%CI 0.84–3.73; $P = 0.002$). However, there was no significant difference in the specificity of CXR with abnormalities suggestive of TB in low compared with high HIV prevalence settings (coefficient –0.23, 95%CI –2.15 to 1.68; $P = 0.812$).

Figure 3 gives the estimates combining sensitivity and 1–specificity of the index test (CXR with abnormalities suggestive of TB) in high and low HIV prevalence settings. The diamonds represent estimates from individual studies in high HIV prevalence settings, the large circle indicates the 95%CIs around the average estimate (small circle crossed by solid line). The squares represent estimates from individual studies in low HIV prevalence settings, and the dotted circle indicates the 95%CIs around the average estimate (small triangle crossed by dotted line.

Prolonged (chronic) cough

Ten studies provided data on 'prolonged cough' (cough of ≥2 weeks) using bacteriologically confirmed TB as the reference standard (Supplementary Table S2). Prolonged cough had very heterogeneous sensitivity, ranging from 14% to 54% (Supplementary Table S2). The pooled DOR was 11.1 (95%CI 7.4–16.5). There was no significant difference in DOR of the included studies with Cochran's *Q* and Higgins's I^2 (11.3 vs. 20.0%; $P = 0.259$). The bivariate diagnostic random-effects meta-analysis indicated that the AUC was 0.78. There was no significant difference in the sensitivity of prolonged cough in low compared with high HIV prevalence settings (coefficient –0.68, 95%CI –1.48 to 0.13; $P = 0.099$). There was also no significant difference in the specificity of prolonged cough in low compared with high HIV prevalence settings (coefficient 0.58, 95%CI –0.50 to 1.67; $P = 0.293$).

Cough of any duration

Seven studies reported on 'cough of any duration'. In general, cough of any duration had higher sensitivity and lower specificity than prolonged cough (Supplementary Table S2). The pooled DOR was 5.5 (95%CI 3.1–9.7). There was no significant difference in DOR of the included studies with Cochran's *Q* and Higgins's I^2 (10.4 vs. 13.8%; $P = 0.316$). The bivariate diagnostic random-effects meta-analysis indicated that the AUC was 0.74. There was no significant difference in the sensitivity of cough of any duration in low compared with high HIV prevalence settings (coefficient 0.02, 95%CI –1.00 to 1.03; $P = 0.977$). There was also no significant difference in the specificity of prolonged cough in low compared with high HIV prevalence settings (coefficient –0.62, 95%CI –2.06 to 0.82; $P = 0.397$).

Any TB symptom

Eleven studies reported on 'any TB symptom' (cough, weight loss, haemoptysis, fatigue, fever/chills, night sweats, difficulty breathing/shortness of breath, pain/tightness in chest, loss of appetite, lymphadenopathy) as a screening tool. The summary sensitivity and specificity by HIV prevalence were not significantly different (Supplementary Table S2 and Supplementary Figure S2).

The pooled DOR was 7.9 (95%CI 4.3–14.3). There was no significant difference in DOR of the included studies with Cochran's *Q* and Higgins's I^2 (9.72 vs. 0%; $P = 0.556$). The bivariate diagnostic random-effects meta-analysis indicated that the AUC was 0.79. There was no significant difference in the sensitivity of any TB symptom in low compared with high HIV prevalence settings (coefficient –0.58, 95%CI –1.50 to 0.34; $P = 0.213$). There was also no significant difference in the specificity of prolonged cough in low compared with high HIV prevalence settings (coefficient 0.43, 95%CI –1.27 to 2.12; $P = 0.620$).

Combination of symptom and chest X-ray

Five studies, all from the previous review, provided data on symptom and/or CXR screening employed in parallel.[6] One study (den Boon et al.[12]) did not provide sufficient data to determine the accuracy of such a screen. Symptom screening plus CXR did not result in the identification of more TB cases (Supplementary Table S2). The addition of cough increased sensitivity by an absolute 3–9% in three studies, but decreased sensitivity by 3% in one study (which used CXR with any abnormality). The specificity decreased in all studies by an absolute 2–5%.[6]

Modelling using any CXR abnormality and any TB symptom as screening tools

We could not obtain data for the sensitivity and specificity of a screening test that combined any TB symptom with any CXR abnormality directly from the studies included in the systematic review. We assumed that persons without symptoms and CXR abnormalities were free from active TB disease. The sensitivity of a screening test that used a combination of CXR and symptoms was therefore assumed to be 100%. In one of the five studies included in the systematic review, which used CXR and symptom screening, 6.5% of the total number of study participants without TB had CXR abnormalities but did not have any TB symptoms. The specificity of the combined screening test (any TB symptom + any CXR abnormality) was thus estimated by deducting 6.5% from the specificity of any TB symptoms (76.5%) (Supplementary Table S3).[21]

If a screening programme is conducted in a setting with a TB prevalence of 1% using any TB symptom (with sensitivity 73.0% and specificity 76.6%), followed by culture (100% sensitivity and 100% specificity), it will have an NPV of 99.7%, with 0.4% of all screened negatives being FN. Conversely, the screening programme will have an NPV of 99.9%, with 0.1% of all screened negatives being FN if any CXR abnormality (with sensitivity 94.1% and specificity 86.8%) is used, followed by culture. If a combination of any TB symptoms and any CXR abnormality is used, the screening test will have an NPV of 100% with a zero FN rate (Supplementary Table S3).

DISCUSSION

Our review found that any TB symptom had the highest sensitivity (73.0%, 95%CI 64.1–80.4), but the lowest specificity (76.5%, 95%CI 61.3–87.0). Prolonged cough had the lowest sensitivity (38.2%, 95%CI 29.1–48.3), but the highest specificity (94.3%, 95%CI 92.2–95.9). Screening for any TB symptom will result in many FP TB suspects, who would require additional confirmatory tests and increase resource needs. The studies using symptom screening, in general, showed considerable heterogeneity in accuracy. Some of this heterogeneity may be explained by the HIV prevalence in the study population. Two studies reported higher sensitivity but lower specificity in HIV-positive participants.[22,23] This may be due to a generally higher prevalence of symptoms and complaints and a faster progression of TB disease in HIV-positive persons than in HIV-negative persons.[24]

Among studies using CXR screening, any CXR abnormality had a higher pooled sensitivity than CXR abnormality suggestive of TB (94.1%, 95%CI 85.8–97.7 vs. 89.3%, 95%CI 81.4–94.1), whereas the pooled specificity was lower in studies using any CXR abnormality than in those using CXR suggestive of TB (86.8%, 95%CI 79.7–91.7 vs. 92.2%, 95%CI 89.7–94.1). The high pooled sensitivity of any CXR abnormality should be interpreted with some caution because there may have been a verification bias, resulting in overdiagnosis and increased sensitivity. In addition, there was no standard definition for any CXR abnormality or CXR abnormality suggestive of TB and, as this may have resulted in inter-reader variability, it is also likely to affect the generalisability of the summary estimates to screening programmes.[13]

We found that the accuracy of 'symptom screening' could be further improved by adding 'CXR screening' to exclude active TB. The addition of CXR increased sensitivity by an absolute 3–9%, whereas the specificity decreased by an absolute 2–5%. Use of any CXR abnormality in patients with any TB symptom will thus have an NPV of 1.00 (Supplementary Table S3). A combination of any TB symptom and any CXR abnormality is thus likely to offer the greatest sensitivity in excluding active TB before LTBI treatment.[25] However, as the specificity of the algorithm was not 100%, it could lead to FP results, which would necessitate further evaluation for TB according to national guidelines.

This review had several limitations that were mainly related to the risk of bias and applicability. Most studies used culture as their reference standard, while two studies used Xpert and three studies used smear as their reference standard. Furthermore, most studies used Löwenstein-Jensen as their reference standard. Although culture is considered to be the 'gold standard', some TB cases might have been missed with culture. Studies that used culture as the reference standard varied in the number of positive culture results (one or two). Reference standards with one sputum sample might result in high sensitivity and a FP rate of the index test because the probability of a positive result from culture increases if more sputum samples are culture tested. The proportion of reference tests conducted varied from 1% to 100%. Most studies applied reference tests only in those with a positive screening result; those with a negative

screening result did not undergo reference tests. This would exclude FN index tests and lead to an overestimation of test sensitivity.[26] Finally, the reference standard was limited to sputum samples. The combination of CXR plus symptom screening would not help identify extra-pulmonary TB, unless it presents with systemic symptoms, such as fever and weight loss.

In conclusion, a screening algorithm using any symptoms of TB and any abnormal CXR findings is likely to offer high sensitivity. This implies that the absence of any TB symptoms and any abnormal CXR findings can be used to exclude active PTB before initiating LTBI treatment among household contacts.

References

1 Houben R M, Dodd P J. The global burden of latent tuberculosis infection: a re-estimation using mathematical modelling. PLOS Med 2016; 13: e1002152.
2 Getahun H, Matteelli A, Chaisson R E, Raviglione M. Latent *Mycobacterium tuberculosis* infection. N Engl J Med 2015; 372: 2127–2135.
3 Rieder H. Epidemiologic basis of tuberculosis control. Paris, France: International Union Against Tuberculosis and Lung Disease, 1999: pp 63–119.
4 World Health Organization. The End TB Strategy. Global strategy and targets for tuberculosis prevention, care and control after 2015. Geneva, Switzerland: WHO, 2014.
5 World Health Organization. Guidelines on the management of latent tuberculosis infection. WHO/HTM/TB/2015.01. Geneva, Switzerland: WHO, 2015.
6 Cameron M, Wakefield M, Trotter L, Inglis G. Exposure to secondhand smoke at work: a survey of members of the Australian Liquor, Hospitality and Miscellaneous Workers Union. Aust NZ J Public Health 2003; 27: 496–501.
7 World Health Organization. Latent tuberculosis infection. Updated and consolidated guidelines for programmatic management. WHO/CDS/TB/2018.4. Geneva, Switzerland: WHO, 2018. http://www.who.int/tb/publications/2018/executivesummary_consolidated_guidelines_ltbi.pdf Accessed April 2019.
8 Hermann K, Van Damme W, Pariyo G W, et al. Community health workers for ART in sub-Saharan Africa: learning from experience—capitalizing on new opportunities. Hum Resour Health 2009; 7: 31.
9 Macaskill P, Gatsonis C, Deeks J, Harbord R, Takwoingi Y. Cochrane handbook for systematic reviews of diagnostic test accuracy. Version 1.0. London, UK: The Cochrane Collaboration, 2010.
10 Sibamo E L, Berheto T M. Community satisfaction with the urban health extension service in South Ethiopia and associated factors. BMC Health Serv Res 2015; 15: 160.
11 Maher D. Clinical features and index of suspicion in adults (HIV-negative and HIV-positive). Tuberculosis: a comprehensive clinical reference. Philadelphia, PA, USA: Saunders, 2009: pp 164–168.
12 Den Boon S, White N, Van Lill S, et al. An evaluation of symptom and chest radiographic screening in tuberculosis prevalence surveys. Int J Tuberc Lung Dis 2006; 10: 876–882.
13 Van't Hoog A, Meme H, 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: 1308–1314.
14 Steingart K R, Henry M, Laal S, et al. Commercial serological antibody detection tests for the diagnosis of pulmonary tuberculosis: a systematic review. PLOS Med 2007; 4: e202.
15 Gatsonis C. Do we need a checklist for reporting the results of diagnostic test evaluations? The STARD proposal. Acad Radiol 2003; 10: 599–600.
16 Whiting P F, Rutjes A W, Westwood M E, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 2011; 155: 529–536.
17 Schünemann H J, Oxman A D, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. Chinese J Evid Based Med 2009; 9: 503–508.
18 Higgins J, Thompson S G. Quantifying heterogeneity in a meta-analysis. Stat Med 2002; 21: 1539–1558.
19 Mathewos B, Owen H, Sitrin D, et al. Community-Based interventions for newborns in Ethiopia (COMBINE): cost-effectiveness analysis. Health Policy Plan 2017; 32 (Suppl 1): i21–i32.
20 Cochrane Centre. RevMan 5.3: user guide. London, UK: Cochrane Centre, 2014.
21 World Health Organization. Tuberculosis prevalence surveys, a handbook. WHO/HTM/TB/2010.17. Geneva, Switzerland: WHO, 2011.
22 Corbett E L, Zezai A, Cheung Y B, et al. Provider-initiated symptom screening for tuberculosis in Zimbabwe: diagnostic value and the effect of HIV status. Bull World Health Organ 2010; 88: 13–21.
23 Ayles H, Schaap A, Nota A, et al. Prevalence of tuberculosis, HIV and respiratory symptoms in two Zambian communities: implications for tuberculosis control in the era of HIV. PLOS ONE 2009; 4: e5602.
24 Corbett E L, Charalambous S, Moloi V M, et al. Human immunodeficiency virus and the prevalence of undiagnosed tuberculosis in African gold miners. Am J Respir Crit Care Med 2004; 170: 673–679.
25 van't Hoog A H, Onozaki I, Lönnroth K. Choosing algorithms for TB screening: a modelling study to compare yield, predictive value and diagnostic burden. BMC Infect Dis 2014; 14: 532.
26 Rutjes A W, Reitsma J B, Di Nisio M, Smidt N, Van Rijn J C, Bossuyt P M. Evidence of bias and variation in diagnostic accuracy studies. CMAJ 2006; 174: 469–476.

RÉSUMÉ

OBJECTIF : Examiner la fiabilité des symptômes, des anomalies de la radiographie pulmonaire (CXR) et de la combinaison des symptômes et de la CXR pour exclure une tuberculose (TB) pulmonaire active avant le traitement d'une infection tuberculeuse latente (LTBI) dans les pays très touchés par la TB.

MÉTHODE : Nous avons mis à jour une revue systématique et une méta analyse afin de calculer les sensibilités, les spécificités, les valeurs prédictives, les odds ratio de diagnostic et la zone sous la courbe pour les tests index. L'analyse a été réalisée selon la méthode de la courbe sommaire d'efficacité du récepteur hiérarchique en R.

RÉSULTATS : Nous avons inclus 24 publications dans la revue systématique et la méta analyse. « Une anomalie quelconque de la CXR » a eu la sensibilité la plus élevée (94,1% (IC95% 85,8–97,7) parmi tous les tests index. « Une anomalie de la CXR suggestive de la TB » a eu une spécificité plus élevée (92,2% ; IC95% 89,7–94,1) qu' « Une anomalie quelconque de la CXR » (86,8% ; IC95% 79,7–91,7). La sensibilité de « Tout symptômes de TB » a été de 73,0% (IC95% 64,1–80,4) tandis qu'une « toux prolongée » (au moins 2 semaines) a eu une spécificité de 94,3% (IC95% 92,2–95,9). Il n'y a pas eu de différence significative de sensibilité et de spécificité dans tous les outils de dépistage stratifiés en fonction du contexte virus de l'immunodéficience humaine (VIH) sauf en cas d'anomalie de la CXR suggestive de TB; il y a eu une sensibilité significativement plus élevée dans les contextes de faible prévalence du VIH, avec des estimations de l'effet à 2,26 (IC95% 0,69–3,82 ; $P = 0{,}002$), que dans les contextes de prévalence élevée du VIH.

CONCLUSION : Dans les pays très affectés par la TB, l'absence d'un « symptôme quelconque de TB » et de « toute anomalie de la CXR » peut être utilisée afin d'exclure une TB pulmonaire active avant la mise en route du traitement d'une LTBI chez des patients âgés de $\geq$5 ans avec des contacts domiciliaires de patients atteints de TB pulmonaire confirmée par bactériologie.

RESUMEN

OBJETIVO: Investigar la precisión de 'índices diagnósticos' definidos por los síntomas, las anomalías de la radiografía de tórax (CXR) o la asociación de síntomas y aspectos radiográficos, con el objeto de descartar la presencia de tuberculosis (TB) pulmonar activa, antes de iniciar el tratamiento de la infección tuberculosa latente (LTBI), en países con alta carga de morbilidad por TB.

MÉTODOS: Se actualizó una revisión sistemática con metanálisis, a fin de calcular las sensibilidades, las especificidades, los valores diagnósticos de un resultado, el cociente de posibilidades de un diagnóstico y el área bajo la curva de eficacia diagnóstica de los 'índices diagnósticos'. En el análisis se aplicó el método de la curva común de eficacia diagnóstica jerárquica en R.

RESULTADOS: Se incluyeron en la revisión sistemática y el metanálisis 24 publicaciones. De todos los índices diagnósticos, la sensibilidad más alta se observó con 'Cualquier anomalía de la CXR' (94,1%; IC95% 85,8–97,7). La 'CXR indicativa de TB' exhibió una especificidad más alta (92,2%; IC95% 89,7–94,1) que 'Cualquier anomalía de la CXR' (86,8%; IC95% 79,7–91,7). La sensibilidad de 'Cualquier síntoma de TB' fue 73,0% (IC95% 64,1–80,4), pero la tos prolongada (durante $\geq$2 semanas) exhibió una especificidad de 94,3% (IC95% 92,2–95,9). No se observó ninguna diferencia significativa de sensibilidad ni especificidad de todos los métodos de tamizaje al estratificarlos según el contexto del VIH, con la excepción de 'Anomalía de la CXR indicativa de TB', que en los entornos con baja prevalencia de infección por el VIH exhibió una sensibilidad significativamente más alta que en los medios con alta prevalencia, con una estimación del efecto de 2,26 (IC95% 0,69–3,82; $P = 0{,}002$).

CONCLUSIÓN: En los países con alta carga de morbilidad por TB se puede utilizar la ausencia de 'Cualquier síntoma de TB' y de 'Cualquier anomalía de la CXR' con el fin de descartar la TB pulmonar activa, antes de iniciar el tratamiento de la LTBI en los niños de $\geq$5 de edad que son contactos domiciliarios de pacientes con TB pulmonar confirmada bacteriológicamente.