Diagnostic accuracy of Xpert MTB/RIF assay for musculoskeletal tuberculosis: a meta-analysis

Hai Wen
Pengzhi Li
Hong Ma
Guohua Lv
Department of Spine Surgery, The Second Xiangya Hospital of Central South University, Changsha, People’s Republic of China

Objectives: Xpert MTB/RIF assay, a rapid and automated real-time nucleic acid amplification test, has been reported for the diagnosis of musculoskeletal tuberculosis (TB) in current years. This meta-analysis aims to determine the diagnostic accuracy of Xpert for the detection of musculoskeletal tuberculosis (TB) and rifampicin resistance.

Methods: We searched PubMed, Embase, China National Knowledge Infrastructure, and Wanfang for original articles published up to 1st June 2017 to identify studies in which the Xpert assay was applied to diagnose musculoskeletal TB. Pooled estimates were calculated using a random-effects model or a fixed-effects model according to heterogeneity. Summary receiver operating characteristic curves and the area under the curve (AUC) were used to summarize overall diagnostic performance. Deeks’ test was performed to evaluate potential publication bias.

Results: Twelve studies were identified with a pooled sensitivity and specificity of respectively 0.81 (95% confidence interval [CI] 0.78–0.83) and 0.83 (95% CI 0.80–0.86) of Xpert for the diagnosis of musculoskeletal TB. Xpert was highly sensitive (0.89, 95% CI 0.79–0.95) and highly specific (0.96, 95% CI 0.92–0.98) in detecting rifampicin resistance. AUC (over 0.9) suggested a relatively high level of overall diagnostic accuracy of Xpert for detecting musculoskeletal TB and rifampicin resistance. Prevalence and reference standard were indicated to be sources of heterogeneity between studies. No publication bias was found.

Conclusion: This study provides available evidence of the rapid and effective role of Xpert in diagnosing musculoskeletal TB and detecting rifampicin resistance.

Keywords: Xpert, diagnostic accuracy, musculoskeletal tuberculosis, meta-analysis, rifampicin resistance

Introduction
Tuberculosis (TB) remains to be one of the predominant infectious diseases and ranks as a leading cause of infection-related mortality worldwide.1,2 Besides the major pulmonary TB (PTB), extrapulmonary sites (EPTB) are often involved during the initial latent phase of the infection.3 EPTB accounts for 10%–42% of TB cases, 10%–25% of which are musculoskeletal TB.4,5 Almost half of all cases of musculoskeletal TB have vertebro involvement.6 Joint destruction, growth arrest, and contractures in large joints, as well as deformity with neurologic compromise in spinal disease, may result in long-term morbidity and disability,7,8 particularly in multidrug-resistant (MDR) cases of TB, which are increasing especially in the developing countries. Therefore, more accurate and timely tests compared with conventional diagnostic approaches are critical for musculoskeletal TB to enable early effective treatment that is associated with reduced deformity and enhanced outcome.
Molecular tests have been investigated previously to identify *Mycobacterium tuberculosis* and mutations conferring resistance to rifampicin (RIF) and/or isoniazid. Xpert MTB/RIF (Xpert/GeneXpert) assay, a rapid and automated real-time nucleic acid amplification test that can detect RIF susceptibility within 2 h, has been endorsed by the World Health Organization (WHO) for the diagnosis of PTB. According to a recent meta-analysis, the pooled sensitivity and specificity of Xpert in diagnosing children’s PTB were 68% and 100%, respectively. The WHO also recommended Xpert for the diagnosis of EPTB in 2013, and this technology was indicated by another meta-analysis to have high specificity but relatively low sensitivity in the detection of EPTB. Applications of Xpert for musculoskeletal TB cases have been reported in current years, but the data are very limited. Held et al reported that Xpert had a sensitivity of 95.6% and a specificity of 96.2% in 71 adult spinal samples compared with culture or histology as a reference standard. Gu et al performed Xpert in 60 cases of bone and joint TB and detected a sensitivity of 82% and a specificity of 100% using composite reference standard (CRS). In consideration of the variable results, we performed this updated meta-analysis to assess the overall diagnostic accuracy of the Xpert assay in the detection of musculoskeletal TB.

**Methods**

**Search strategy and study selection**

This meta-analysis was conducted according to the criteria of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses group. A systematic search was performed using PubMed, Embase, China National Knowledge Infrastructure, and Wanfang for original articles published up to 1st June 2017 to identify studies in which the Xpert assay was applied to diagnose musculoskeletal TB. The following search terms were used: “Xpert” OR “GeneXpert” AND “tuberculosis” OR “TB” AND “bone” OR “joint” OR “spine”. There were no language restrictions. References within selected articles were also reviewed to identify additional relevant studies.

Titles and abstracts of all studies were independently screened, and afterwards, full-text publications of relevant articles were read by two investigators (H Wen and P Li). Disagreements were resolved by consensus. A study was included when it met the following inclusion criteria: an original research article written in English or Chinese; the Xpert assay was used for the diagnosis of musculoskeletal TB, and the reference standard was described; and sufficient data including true-positive, false-positive, false-negative, and true-negative results of the Xpert assay were provided.

**Data extraction and quality assessment**

The same two reviewers independently extracted data from each included study, and discrepancies were resolved by consensus. The retrieved data included author, publication year, country, sample size, specimen types, reference standard, and numbers of true-positive, false-positive, false-negative, and true-negative results. The methodological quality of each study was assessed using the validated Quality Assessment of Diagnostic Accuracy Studies (QUADAS) tool.

**Statistical analysis**

We used Meta-DiSc 1.4 (XI Cochrane Colloquium, Barcelona, Spain) to estimate the following data of individual study: sensitivity, specificity, positive likelihood ratio (PLR), negative likelihood ratio (NLR), and diagnostic odds ratio (DOR) together with 95% confidence intervals (95% CIs). A summary receiver operating characteristic (SROC) curve was plotted on the basis of pooled sensitivity and specificity, and the overall diagnostic performance of Xpert was displayed by the area under the curve (AUC). Heterogeneity between studies was evaluated using the Q test and I² statistics, according to which pooled indices across studies were calculated using a random-effects model or a fixed-effects model. Deeks’ funnel plot was generated using STATA 12.0 (Stata Corp., College Station, TX, USA) to detect the potential publication bias. All statistical tests were two-sided, and *p* values less than 0.05 were considered statistically significant.

**Results**

**Search result and study characteristics**

Our search yielded a total of 466 citations, 12 of which were finally included in the current meta-analysis based on the inclusion criteria (Figure 1 and Table 1). The included studies were published between 2014 and 2017. Data came from the People’s Republic of China, South Africa, India, and Indonesia, all of which are countries with high TB burden and low/middle income. The studies totally enrolled 1690 specimens with an average sample size of 141 (range, 49–338). All the studies described the performance of the Xpert assay in suspected musculoskeletal TB samples which were tissue or pus taken from bones, joints, or spine, but the individual reference standard to assess the diagnostic accuracy of Xpert was not consistent, including mycobacterial culture only (141 specimens), culture or histology (586 specimens), and CRS (963 specimens). The CRS included all patients who have...
improved both clinically and radiologically after empirical antituberculous therapy where the results of culture, smear, and histology might be positive or negative for TB.²⁰ Moreover, the mean score of the 12 studies was 9.42 (range, 9–11) according to the criteria of QUADAS (Table 1).

Diagnostic accuracy of Xpert for musculoskeletal TB detection

Pooled sensitivity and specificity of Xpert in musculoskeletal TB detection are shown in Figure 2A. Among the 12 included studies, sensitivity ranged from 0.69 to 1.00 (pooled 0.81, 95% CI 0.78–0.83), and specificity ranged from 0.17 to 1.00 (pooled 0.83, 95% CI 0.80–0.86). Additionally, pooled parameters including PLR (10.90, 95% CI 3.11–38.20), NLR (0.22, 95% CI 0.19–0.25), and DOR (50.86, 95% CI 22.00–117.56) were calculated. Reflected by $I^2$ statistics of the indices (56.5% for sensitivity and 94.2% for specificity), there existed heterogeneity in diagnostic performance between studies. An SROC curve displaying sensitivity against 1 − specificity from individual study was plotted for Xpert in musculoskeletal TB, with an AUC of 0.9182 (Figure 3A), suggesting high overall diagnostic accuracy of the Xpert assay in detection of musculoskeletal TB.

Diagnostic accuracy of Xpert in detection of RIF-resistant musculoskeletal TB

There were five studies which reported the results of RIF resistance in culture-positive musculoskeletal TB clearly...
Among them, one study\(^{23}\) was not included in the analysis due to no positive results of Xpert for detection of RIF-resistant musculoskeletal TB. The pooled sensitivity (0.89, 95% CI 0.79–0.95) and specificity (0.96, 95% CI 0.92–0.98) in diagnosing RIF-resistant musculoskeletal TB for four studies are displayed in Figure 2B. The pooled PLR, NLR, and DOR were 16.00 (95% CI 2.98–86.04), 0.13 (95% CI 0.06–0.25), and 128.60 (95% CI 13.69–1208.09), respectively. The SROC curve is illustrated in Figure 3B, and the AUC was 0.9264.

**Heterogeneity and subgroup analysis**  
Indicated by the Spearman correlation coefficient (0.483, \(p = 0.112\)), heterogeneity between studies was not correlated.
with threshold effect. We hypothesized that the most likely factors causing heterogeneity were as follows: sample size, prevalence, study design, specimen type, reference standard, and QUADAS scores. Heterogeneity analysis showed that prevalence and reference standard were the probable sources of heterogeneity (both p < 0.05). Therefore, we stratified the 12 studies into subgroups by prevalence and reference standard. As shown in Table 2, the sensitivity and specificity of Xpert in studies with lower prevalence (<50%) respectively elevated and declined compared with those with higher prevalence (≥50%). Besides, Xpert was most sensitive (0.96, 95% CI 0.86–1.00) and least specific (0.47, 95% CI 0.36–0.57) in studies using culture as the reference standard. In contrast, Xpert was most specific (0.97, 95% CI 0.94–0.99) and least sensitive (0.79, 95% CI 0.76–0.82) in studies using CRS as the reference standard. Reflected by I² (46.0% for sensitivity and 22.4% for specificity), additionally, there was no significant heterogeneity of the diagnostic accuracy of Xpert in studies using CRS as the reference standard.

**Table 2** Subgroup analysis of the results of Xpert for detecting musculoskeletal TB

<table>
<thead>
<tr>
<th>Variables</th>
<th>Number of studies</th>
<th>Sensitivity (95% CI)</th>
<th>I²</th>
<th>Specificity (95% CI)</th>
<th>I²</th>
<th>AUC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence (%)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥50</td>
<td>9</td>
<td>0.80 (0.77–0.82)</td>
<td>32.2%</td>
<td>0.88 (0.84–0.91)</td>
<td>83.5%</td>
<td>0.8820</td>
</tr>
<tr>
<td>&lt;50</td>
<td>3</td>
<td>0.89 (0.79–0.95)</td>
<td>79.1%</td>
<td>0.72 (0.65–0.79)</td>
<td>98.4%</td>
<td>0.9568</td>
</tr>
<tr>
<td>Reference standard</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culture only</td>
<td>2</td>
<td>0.96 (0.86–1.00)</td>
<td>59.2%</td>
<td>0.47 (0.36–0.57)</td>
<td>97.5%</td>
<td>NE</td>
</tr>
<tr>
<td>Culture/histology</td>
<td>3</td>
<td>0.81 (0.76–0.85)</td>
<td>0.0%</td>
<td>0.85 (0.80–0.89)</td>
<td>94.7%</td>
<td>0.8780</td>
</tr>
<tr>
<td>CRS</td>
<td>7</td>
<td>0.79 (0.76–0.82)</td>
<td>46.0%</td>
<td>0.97 (0.94–0.99)</td>
<td>22.4%</td>
<td>0.9429</td>
</tr>
</tbody>
</table>

**Publication bias**

Deeks’ funnel plot asymmetry test was performed to evaluate potential publication bias of included studies. In spite of some asymmetry shown in the funnel plots due to the limited quantity of studies (Figure 4), the p value (0.621) of Deeks’ test was not statistically significant, indicating low likelihood of publication bias.

**Discussion**

Musculoskeletal TB is the third most common type of EPTB after pleural and lymphatic disease, and the spine and weight-bearing joints are the most vulnerable sites of infection. Although identification of *M. tuberculosis* and drug susceptibility tests remain the gold standard for TB diagnosis, pathogen culturing is time-consuming and owns a relatively high false-negative rate, together with the increasing incidence of MDR-TB, leading to a pressing need for more timely and effective diagnostic methods. A rising and promising rapid test, Xpert MTB/RIF assay, has been

![Figure 4](https://example.com/fig4.png)

**Figure 4** Deeks’ funnel plot indicating no significant risk of publication bias (p = 0.621). Solid circles represent each study included in the meta-analysis.

**Abbreviation:** ESS, effective sample size.
identified to diagnose TB and detect RIF resistance rapidly and effectively.\textsuperscript{9,10} This assay was reported to detect as low as 131 colony forming units per milliliter of \textit{M. tuberculosis} culture from a specimen.\textsuperscript{81} In the current study, we evaluated the diagnostic accuracy of Xpert for the detection of musculoskeletal TB and RIF resistance.

Our data showed that the pooled sensitivity and specificity of Xpert in detecting musculoskeletal TB were 81\% and 83\%, respectively. The efficacy of Xpert is lower (sensitivity of 89\% and specificity of 99\%) in the diagnosis of PTB according to a previous meta-analysis,\textsuperscript{32} which may be owing to less content of \textit{M. tuberculosis} in musculoskeletal specimens compared with sputum from patients with PTB.\textsuperscript{20} Penz et al\textsuperscript{14} described a pooled sensitivity and specificity of respectively 77\% and 97\% of Xpert in the detection of EPTB in a meta-analysis including 36 studies. Relatively homogeneous tested samples and less included studies, especially some describing comparatively low specificity,\textsuperscript{25,26} may account for the reduced pooled specificity in our study. In addition, a pooled PLR of 1.090 and a pooled NLR of 0.22 indicate that patients with positive or negative Xpert results have a 10.90-fold higher or a 0.22-fold lower chance of suffering from musculoskeletal TB than not. These data may not be convincing enough for clinical use. Nevertheless, a pooled DOR of 50.86 indicates significant discriminant effect, and an AUC of 0.9182 suggests a relatively high level of overall diagnostic accuracy of Xpert for detecting musculoskeletal TB. It is generally known that simultaneous assessment of RIF resistance is a crucial advantage of Xpert. This test was revealed by our analysis to have high pooled sensitivity (89\%) and specificity (96\%) in the diagnosis of RIF-resistant musculoskeletal TB, which is consistent with the high efficacy of Xpert in detecting RIF-resistant PTB reported by Chang et al.\textsuperscript{33} Moreover, other parameters including pooled PLR (16.00), pooled NLR (0.13), pooled DOR (128.60), and AUC (0.9264) support that Xpert meets the requirements of effective detection of RIF resistance.

Expectedly, significant heterogeneity of results between studies was observed, whereas it was not correlated with threshold effect. Due to the limited data of included studies in this meta-analysis, we restricted meta-regression analysis to six factors mentioned above, and attributed the heterogeneity to prevalence and reference standard. It was indicated by subgroup analysis that in studies with lower prevalence of musculoskeletal TB, Xpert showed increased sensitivity and reduced specificity, which could be explained by relatively less positive and more negative results detected by the reference standard used in these studies. Also, diagnostic accuracy varied significantly using different reference standards. Although mycobacterial culture has been widely accepted as a gold standard for diagnosing TB, its relatively high false-negative rate may raise negative results and decrease positive results. Hence, high sensitivity (96\%) and low specificity (47\%) of Xpert were pooled when using this gold standard only. We meanwhile observed a much higher pooled specificity (97\%) of Xpert when CRS was used as the reference standard, which was probably because of reduced false-negative results according to CRS.

Several limitations of the current meta-analysis should be considered. Lack of included studies, particularly those recording the diagnostic performance of Xpert for RIF resistance, may limit the statistical power for drawing definite conclusions about the efficacy of Xpert in detecting musculoskeletal TB and RIF-resistant cases. Also, three included studies\textsuperscript{24,28,29} designed as case–control probably overvalued the diagnostic accuracy of Xpert as a result of diagnostic suspicion bias. In addition, most studies did not mention the information on blinding, and sample collection and processing were variable across the studies. The pooled estimates should be interpreted with caution in consideration of existing heterogeneity among the studies.

In conclusion, our study demonstrates that Xpert MTB/RIF assay has high sensitivity and specificity in the diagnosis of musculoskeletal TB and detection of RIF resistance, providing useful information for clinical decision-making. More efforts should be made to define the accuracy of this promising test for diagnosing musculoskeletal TB in the future.

Acknowledgments
The authors acknowledge the grants from the Fundamental Research Funds for the Central Universities of Central South University (No. 2016zzts149).

Disclosure
The authors report no conflicts of interest in this work.

References


