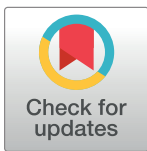RESEARCH ARTICLE

Diagnostic accuracy of the Xpert MTB/RIF assay for bone and joint tuberculosis: A meta-analysis

Yanqin Shen[☯], Guocan Yu[☯], Fangming Zhong, Xiaohua Kong[iD]*

Zhejiang Tuberculosis Diagnosis and Treatment Center, Zhejiang Chinese Medicine and Western Medicine Integrated Hospital, Hangzhou, Zhejiang, China

☯ These authors contributed equally to this work.
* kongxiaohuamed@sina.com

Abstract

Background

This study aimed to evaluate the accuracy of the Xpert MTB/RIF assay for the diagnosis of bone and joint tuberculosis.

Methods

We searched databases from their inception to May 7, 2019 for published articles and reviewed them to assess the accuracy of Xpert MTB/RIF with respect to a composite reference standard (CRS) and mycobacterial culture. Meta-analyses were performed using a bivariate random-effects model, and the sources of heterogeneity were assessed via subgroup analysis and meta-regression.

Results

Nineteen independent (9 prospective, 5 retrospective, and 5 case-control) studies that compared Xpert MTB/RIF with the CRS and 14 (6 prospective, 7 retrospective, and 1 case-control) studies that compared it with culture were included. The pooled sensitivity and specificity of Xpert MTB/RIF were 81% (95% confidence interval [CI], 77–84) and 99% (95% CI, 97–100) compared to the CRS, respectively, and 96% (95% CI, 90–98) and 85% (95% CI, 57–96) compared to culture, respectively. The pooled sensitivity and specificity using pus samples vs. the CRS were 82% (95% CI, 76–86) and 99% (95% CI, 95–100), respectively. The proportions obtained while working with tissue samples vs. the CRS were 84% (95% CI, 76–90) and 98% (95% CI, 94–99), respectively. There was no significant difference in diagnostic accuracy among the types of specimens.

Conclusions

Xpert MTB/RIF demonstrates good diagnostic accuracy for bone and joint tuberculosis, the results of which are not related to the type of specimen.

Introduction

Tuberculosis is a major infectious disease globally and poses a serious threat to public health [1]. Extrapulmonary tuberculosis (EPTB) accounts for about 10% of all tuberculosis cases [1]. Bone and joint tuberculosis (BJTB) is a common type of EPTB and accounts for about 10–34% of EPTB cases [2, 3]. BJTB can lead to joint destruction, deformity, and even paraplegia, thereby seriously affecting the quality of life. Therefore, early and correct diagnosis and treatment are critical [4]. However, early diagnosis is very difficult due to the atypical symptoms of BJTB, deep lesions, and difficulty in obtaining specimens [5]. Traditional diagnostic protocols, such as *Mycobacterium tuberculosis* culture, are quite time-consuming and have a low sensitivity [6]. Therefore, rapid laboratory diagnosis of BJTB is an urgent necessity. The Xpert MTB/RIF assay is a rapid, automated molecular test with a high accuracy for the detection of pulmonary tuberculosis (PTB) and EPTB [7]. This assay has also been recommended for the diagnosis of lymph node tuberculosis and has shown good diagnostic accuracy [8]. However, the diagnostic accuracy of this assay for BJTB remains controversial. Due to the lack of independent systematic research on the diagnostic accuracy of Xpert MTB/RIF assay for BJTB, the possible influence of the type of specimen (pus and tissue samples) on the results is yet to be clarified. Therefore, we performed a meta-analysis to confirm the diagnostic accuracy of the Xpert MTB/RIF assay, compared to that of the composite reference standard (CRS) and mycobacterial culture, in the detection of BJTB. We assessed the pooled sensitivity and specificity of this assay compared to different references. Moreover, the diagnostic accuracy of the test was evaluated, based on different sample types, lesion sites, conditions of samples, and patient selection methods by subgroup analysis.

Materials and methods

Data sources and search strategy

We searched PubMed, Embase, the Cochrane Library, the Wanfang database, and China National Knowledge Infrastructure for studies evaluating the diagnostic accuracy of Xpert for BJTB on May 7, 2019. The search formula ((Xpert OR Gene Xpert) AND ("Tuberculosis, Osteoarticular"[Mesh] OR "Tuberculosis, Spinal"[Mesh] OR "Extra pulmonary tuberculosis")) was used for PubMed without any limitation. Similar search formulae were used for Embase, the Cochrane Library, China National Knowledge Infrastructure, and Wanfang databases. References cited in the included articles and reviews were further explored for possible candidate studies.

Inclusion criteria

We included full-text original studies that assessed the diagnostic accuracy of Xpert assay for BJTB using bone and joint specimens. Reference standards were well-defined and appropriate to the studies. The articles directly provided true positive (TP), false positive (FP), false negative (FN), and true negative (TN) values for the assay or included the data necessary to calculate these measures. We excluded case reports, articles written in languages other than English and Chinese, studies with < 10 samples, conference reports, and abstracts without full articles.

Reference standard

A composite reference standard (CRS) or mycobacterial culture was defined as the reference standard in our study. Clinical symptoms, radiographic features, biochemical test results, smears, culture, histopathology, and response to anti-tuberculosis drugs constituted the reference standards in the CRS. Some or all of the factors with positive results were considered

positive for BJTB. Cases were considered as non-BJTB if all the results were negative. We used the CRS as defined in the original paper.

Literature screening and selection

Two investigators (Guocan Yu and Yanqin Shen) independently assessed the candidate articles by reviewing their titles and abstracts, followed by the full text, for inclusion. Discrepancies between the two investigators were resolved by discussion with a third investigator (Xiaohua Kong).

Data extraction

We extracted data including author name; year; country; TP, FP, FN, and TN values for the assay; reference standard; patient selection method; some steps (e.g., homogenization); specimen type; and condition along with other parameters. The same two investigators independently extracted the necessary information from each of the included articles; we cross-checked the information they obtained. Discrepancies in the two data sets were settled by a discussion with a third investigator, similar to that used during the literature selection phase. Data from studies against two different reference standards were treated separately.

Assessment of study quality

Based on the two reference standards (CRS and culture), the two investigators independently divided the studies into two groups and used a revised tool for Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) to assess study quality separately [9]. We chose not to carry out formal assessment of publication bias, as the available methods such as funnel plots are not considered valid for diagnostic accuracy reviews [10].

Data synthesis and statistical analysis

We first obtained the values corresponding to TP, FP, FN, and TN in each included study, and calculated the estimated pooled sensitivity and specificity of Xpert MTB/RIF associated with the 95% confidence interval (CI), against CRS or culture, using bivariate random-effects models. Forest plots for sensitivity and specificity were generated for each study. The areas under summary receiver operating characteristic (SROC) curves (AUC) were subsequently calculated. I^2 statistics were used to assess heterogeneity between the studies and a reference standard. While 0% indicated no observed heterogeneity, values greater than 50% were considered to imply substantial heterogeneity [11]. We explored different types of samples, different lesion sites, different patient selection method, decontamination methods, sample conditions, and homogenization as potential sources of heterogeneity, using subgroup and meta-regression analyses. At least four published studies were required to perform the meta-analysis for predefined variable types. Data from studies against CRS and culture were analyzed separately. Stata version 14.0 (Stata Corp., College Station, TX, USA) with the *midas* command packages was used to generate forest plots of sensitivity and specificity with 95% CI for each study and carry out meta-analyses and meta-regression analyses.

Imperfect reference standards

Imperfect reference standards may lead to misclassification of samples in diagnostic validity studies [12, 13]. Due to the paucibacillary nature of EPTB, a culture would be an imperfect reference standard and lead to an underestimation of the true specificity of Xpert MTB/RIF. A CRS is a composite standard that comprises results from several tests; however, a CRS itself

may have reduced specificity, thereby leading to apparent FN Xpert MTB/RIF results, an underestimation of the true sensitivity of Xpert MTB/RIF [12, 14]. Therefore, a study comparing Xpert MTB/RIF with both culture and CRS might provide a more credible range for sensitivity and specificity.

Results

Identification of studies and study characteristics

Through our search strategy, four hundred candidate articles were identified from relevant databases, and twenty-six qualified articles were included according to the inclusion criteria (Fig 1) [15–40], including 12 prospective studies, 9 retrospective studies and 5 case-control studies. The kappa index of agreement for the selection and data extraction was 0.877 (95% CI, 0.779–0.975) between the two investigators. Only two studies were conducted in high-income countries [18, 35], while the rest were conducted in low- and middle-income countries. Fifteen articles were written in English and eleven in Chinese. A median of 109 specimens was evaluated in each article (range, 13–418). We excluded five other articles that reported on the sensitivity only, without reporting any specificity [41–45]. Two other articles published in languages other than English or Chinese were excluded [46, 47]. The specimens used in the studies included pus, tissue, and joint fluid. The commonest site of infection was the spine.

Articles that reported the use of two different reference standards in the same study were considered to include two independent studies. In accordance with this principle, 33 independent studies were included: 19 (9 prospective studies, 5 retrospective studies, and 5 case-control studies) compared Xpert MTB/RIF with CRS, and 14 (6 prospective studies, 7 retrospective studies, and 1 case-control study) compared Xpert MTB/RIF with culture (Table 1).
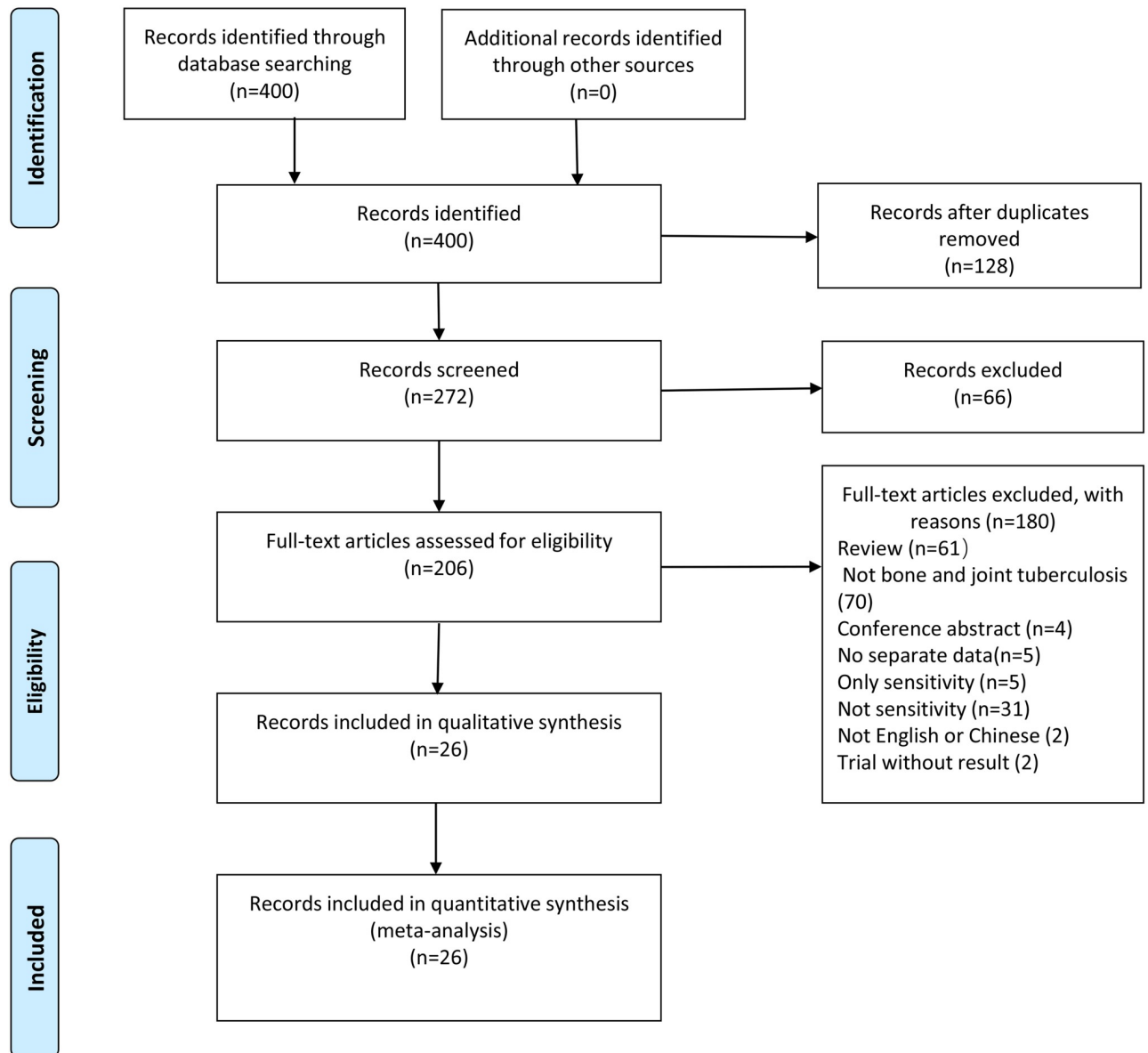
Study quality

The overall methodological quality of the included studies, using a CRS and culture, is summarized in Fig 2. The risk of bias was mainly due to patient selection and the reference standard. The flow and timing of the risk of bias from the index test was judged to be relatively low.

Diagnostic accuracy of Xpert MTB/RIF assay for BJTB

Nineteen studies included a comparison of 3173 samples with a CRS. The sensitivity of Xpert MTB/RIF ranged from 67% (95% CI, 49–81) to 100% (95% CI, 29–100). The pooled sensitivity of Xpert MTB/RIF assay for BJTB was 81% (I^2 = 72%; 95% CI, 77–84) and the specificity ranged from 90% (95% CI, 68–99) to 100% (95% CI, 99–100). The pooled specificity of Xpert MTB/RIF was 99% (I^2 = 63%; 95% CI, 97–100) (Fig 3). There was substantial heterogeneity of sensitivity and specificity. The AUC of the SROC was 0.93 (95% CI, 0.91–0.95).

When compared to a culture standard, 14 studies with 1884 samples were included. The sensitivity of Xpert MTB/RIF ranged from 50% (95% CI, 12–88) to 100% (95% CI, 88–100). The pooled sensitivity of Xpert MTB/RIF was 96% (I^2 = 91%; 95% CI, 90–98). The specificity of Xpert MTB/RIF ranged from 17% (95% CI, 9–28) to 100% (95% CI, 99–100) and the pooled specificity was 85% (I^2 = 97%; 95% CI, 57–96) (Fig 4). As expected, the sensitivity was improved, and the specificity was undervalued when compared to the culture. There was substantial heterogeneity of sensitivity and specificity. The AUC of the SROC was 0.97 (95% CI, 0.96–0.98) vs. that of the culture, suggesting an excellent overall diagnostic accuracy.

We explored the heterogeneity among studies, using subgroup and meta-regression analyses on predefined subgroups of sample types, lesion sites, patient selection methods, decontamination methods, sample conditions, and homogenization methods used in the assay. The

Identification

Records identified through database searching (n=400)

Additional records identified through other sources (n=0)

Records identified (n=400)

Records after duplicates removed (n=128)

Screening

Records screened (n=272)

Records excluded (n=66)

Eligibility

Full-text articles assessed for eligibility (n=206)

Full-text articles excluded, with reasons (n=180)
Review (n=61)
Not bone and joint tuberculosis (70)
Conference abstract (n=4)
No separate data(n=5)
Only sensitivity (n=5)
Not sensitivity (n=31)
Not English or Chinese (2)
Trial without result (2)

Records included in qualitative synthesis (n=26)

Included

Records included in quantitative synthesis (meta-analysis) (n=26)

Fig 1. Literature retrieval flow chart. 124, 22, 228, 15, 11 articles were found from Pubmed, the Cochrane Library, Embase, Wanfang database and CNKI respectively.

https://doi.org/10.1371/journal.pone.0221427.g001

sensitivity of Xpert MTB/RIF using pus samples ranged from 71% (95% CI, 60–80) to 94% (95% CI, 83–99) and the specificity ranged from 94% (95% CI, 80–99) to 100% (95% CI, 96–100) compared with the CRS. The pooled sensitivity was 82% (I^2 = 68%; 95% CI, 76–86), and the pooled specificity was 99% (I^2 = 8%; 95% CI, 95–100) (Fig 5a). There was an acceptable level of heterogeneity in the specificity and a substantial level of heterogeneity in the sensitivity among studies of Xpert MTB/RIF assay using pus samples compared to the CRS. The AUC of the SROC was 0.95 (95% CI, 0.93–0.97), suggesting a very good overall diagnostic accuracy. Using tissue samples, the sensitivity ranged from 74% (95% CI, 52–90) to 96% (95% CI, 85–99), and the specificity ranged from 92% (95% CI, 81–98) to 100% (95% CI, 96–100) compared with the CRS. The pooled sensitivity and specificity of Xpert MTB/RIF assay using tissue

Table 1. Characteristics of the included studies.

Author	Year	County	Sample type	Reference	TP	FP	FN	TN	Decontaminate method	Sample condition	Homogenisation	SR	Patient selection method
Held, M.a	2014	South Africa	spinal tissue	culture	25	9	2	35	No	fresh	Mechanical	3:1	consecutive
Held, M.b	2014	South Africa	spinal tissue	crs	44	1	2	25	No	fresh	Mechanical	3:1	consecutive
Litao Li	2014	China	spinal tissue or pus	culture	63	26	1	50	NALC-NaOH	frozen	Mechanical	-	Convenience
Gu,Y.a	2015	China	bone and joint pus	crs	41	0	9	10	No	fresh	No	2:1	Convenience
Gu,Y.b	2015	China	bone and joint pus	culture	24	17	0	19	No	fresh	No	2:1	Convenience
Kim, Y. W.a	2015	South Korea	joint fluid	culture	3	0	0	280	NALC-NaOH	frozen	No	-	consecutive
Kim, Y. W.b	2015	South Korea	joint fluid	crs	3	0	0	280	NALC-NaOH	frozen	No	-	consecutive
Hongmei Chen	2015	China	bone and joint pus	crs	100	0	30	9	No	fresh	No	-	Convenience
Wenyun Jia	2015	China	bone and joint pus	crs	46	1	3	31	No	fresh	No	2:1	Convenience
Held, M.	2016	South Africa	bone and joint tissue	crs	17	0	6	86	No	fresh	No	3:1	consecutive
Zhen Li.a	2016	China	bone and joint tissue or pus	crs	199	0	54	59	NALC-NaOH	fresh	Mechanical	2:1	Convenience
Zhen Li.b	2016	China	bone and joint tissue or pus	culture	121	78	1	112	NALC-NaOH	fresh	Mechanical	2:1	Convenience
Kai Tang	2016	China	knee joint fluid	crs	20	2	9	18	No	fresh	No	2:1	consecutive
Arockiaraj, J. a	2017	India	spinal tissue or pus	crs	180	0	73	85	NALC-NaOH	fresh	No	-	Convenience
Arockiaraj, J. b	2017	India	spinal tissue or pus	culture	99	82	13	144	NALC-NaOH	fresh	No	-	Convenience
Held, M.	2017	South Africa	bone and joint tissue	crs	84	1	7	114	No	fresh	No	3:1	consecutive
Jin, Y. H.	2017	China	spinal pus	crs	63	0	18	28	No	fresh	Mechanical	2:1	Convenience
Li, Y.	2017	China	joint fluid	culture	26	2	1	6	No	fresh	No	2:1	Convenience
Massi, M. N.	2017	Indonesia	vertebral bone tissues.	culture	22	40	0	8	NALC-NaOH	fresh	No	2:1	Convenience
Tang, L.	2017	China	spinal tissue or pus	culture	96	57	3	67	NALC-NaOH	frozen	No	3:1	Convenience
Weijie Dong	2017	China	bone and joint tissue	crs	196	0	55	21	NALC-NaOH	fresh	Mechanical	2:1	Convenience
Zheng Zhou	2017	China	spinal tissue	crs	55	1	17	19	No	fresh	Mechanical	2:1	Convenience
Zhujun Zhu	2017	China	bone and joint pus	crs	55	1	12	29	No	fresh	Mechanical	2:1	Convenience
Khan, A. S.a	2018	Pakistan	joint fluid	culture	6	0	1	6	No	fresh	No	-	Convenience
Khan, A. S.b	2018	Pakistan	Bone marrow	culture	3	0	3	30	No	fresh	No	-	Convenience
Li, Y.	2018	China	bone and joint pus	crs	24	0	12	13	NALC-NaOH	fresh	Mechanical	2:1	consecutive
Perez-Risco, D.	2018	Spain	bone and joint tissue or pus	culture	14	0	3	23	No	frozen	No	2:1	Convenience
Tang, Y.	2018	China	joint pus	crs	61	0	25	30	No	fresh	No		Convenience
Wang, G.	2018	China	spinal pus	crs	272	0	47	99	NALC-NaOH	fresh	No	2:1	Convenience
Tingting Hou	2018	China	spinal tissue	crs	155	4	37	46	NALC-NaOH	fresh	Mechanical	2:1.5	Convenience

(Continued)

Table 1. (Continued)

Author	Year	County	Sample type	Reference	TP	FP	FN	TN	Decontaminate method	Sample condition	Homogenisation	SR	Patient selection method
Chenguang Jia	2018	China	spinal pus	culture	69	8	1	90	-	frozen	No	1.5:1	Convenience
Hongwei Liu.a	2018	China	bone and joint pus	crs	84	2	11	32	No	fresh	Mechanical	2:1	Convenience
Hongwei Liu.b	2018	China	bone and joint pus	culture	30	54	0	11	No	fresh	Mechanical	2:1	Convenience

CRS, composite reference standard; TP, true positive; FP, false positive; FN, false negative; TN, true negative; SR, sample ratio

samples vs. CRS were 84% (I^2 = 79%; 95% CI, 76–90) and 98% (I^2 = 57; 95% CI, 94–99), respectively (Fig 5b). There was a substantial level of heterogeneity among studies of Xpert MTB/RIF assay using tissue samples compared to the CRS. The AUC of the SROC was 0.97 (95% CI, 0.95–0.98). Compared to the CRS, studies with pus and tissue samples showed similar levels of sensitivity (82% vs. 84%, P > 0.05) and specificity (98% vs. 99%, P > 0.05), respectively. Only two studies reported separate data regarding joint fluid vs. CRS, and meta-analysis of this specimen type was not performed. When culture results were used as the reference standard, separate data using pus, tissue, and joint fluid were limited, and meta-analysis could not be performed.

For spinal tuberculosis, the sensitivity ranged from 71% (95% CI, 65–77) to 96% (95% CI, 85–99), and the specificity ranged from 92% (95% CI, 81–98) to 100% (95% CI, 96–100)
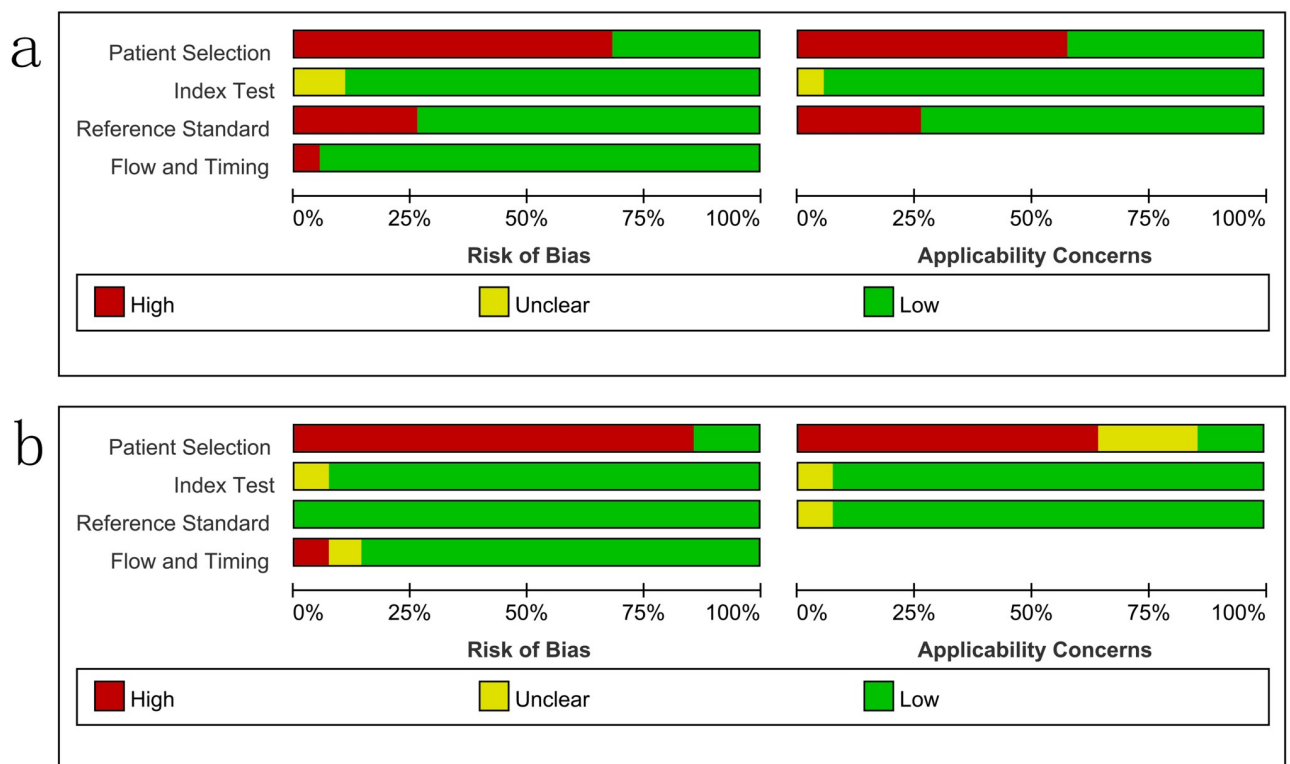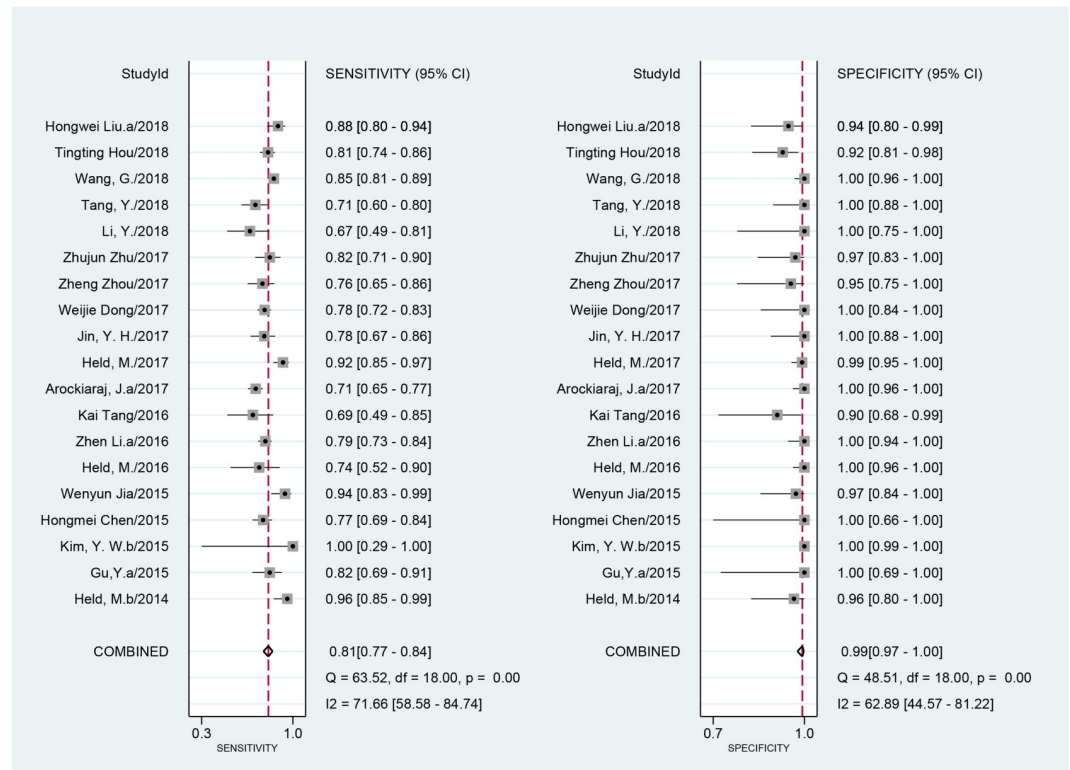

Fig 2. Methodological quality graphs (risk of bias and applicability concerns) as percentages across the included studies. a: composite reference standard. b: culture reference standard.
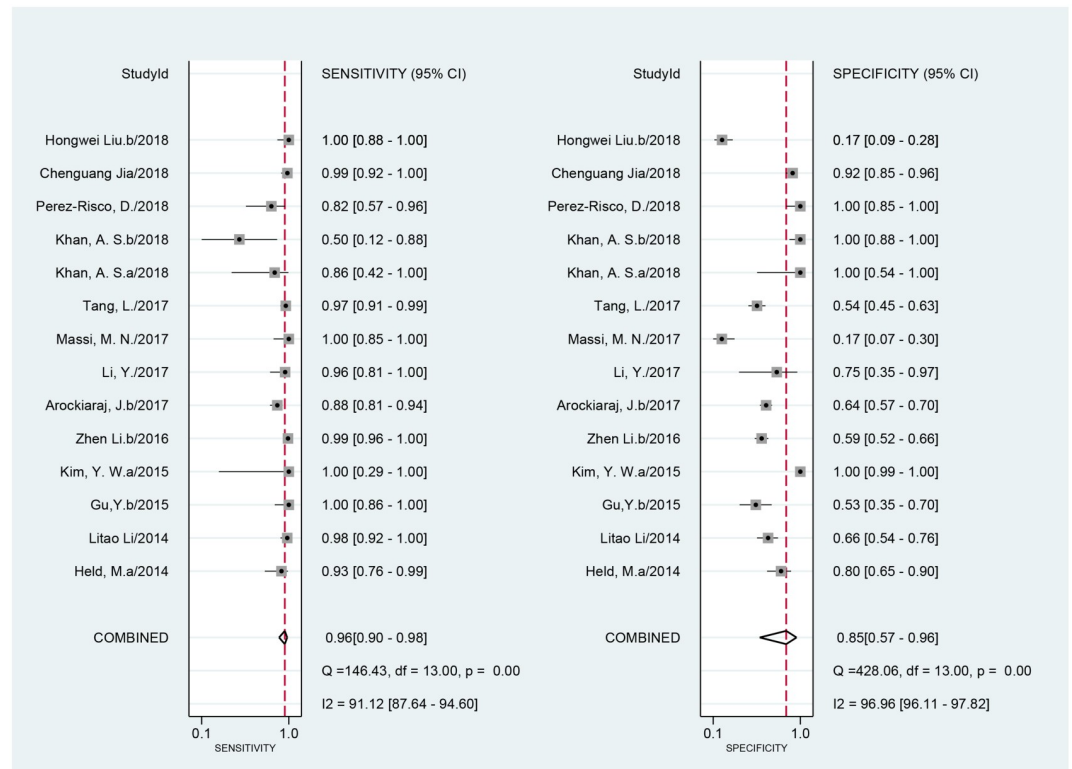
Fig 3. Forest plot of Xpert sensitivity and specificity for tuberculosis detection in BJTB compared with a composite reference standard. BJTB: bone and joint tuberculosis.

compared with the CRS. The pooled sensitivity and specificity of Xpert MTB/RIF assay for spinal tuberculosis against the CRS were 81% (I^2 = 86%; 95% CI, 75–86) and 99% (I^2 = 69%; 95% CI, 93–100), respectively (Fig 6a). There was a substantial level of heterogeneity among studies. The AUC of the SROC was 0.93 (95% CI, 0.90–0.95). When compared to culture, the sensitivity ranged from 88% (95% CI, 81–94) to 100% (95% CI, 85–100), and the specificity ranged from 17% (95% CI, 7–30) to 92% (95% CI, 85–96). The pooled sensitivity and specificity of Xpert MTB/RIF assay for spinal tuberculosis against culture were 97% (I^2 = 74%; 95% CI, 91–99) and 64% (I^2 = 95%; 95% CI, 41–83) (Fig 6b). The heterogeneity among studies was also substantial. The AUC of the SROC was 0.96 (95% CI, 0.94–0.97). A meta-analysis could not be performed due to limited data on other bone tuberculosis sites.

Meta-regression analysis showed that patient selection method (convenience or consecutive) might affect the sensitivity of the Xpert MTB/RIF assay (80% vs. 83%, meta-regression P < 0.01) but had no effect on its specificity (99% vs. 99%, meta-regression P > 0.05) when compared to the CRS. When compared to culture, the patient selection method did not have any effect on the sensitivity and specificity of the Xpert MTB/RIF assay (meta-regression P > 0.05). Meta-regression analysis showed that decontamination method (with or without N-acetyl-L-cysteine / sodium hydroxide), sample condition (fresh or frozen), method of homogenization (mechanical or otherwise) did not have any effect on the sensitivity and specificity of the Xpert MTB/RIF assay (meta-regression P > 0.05), compared to that of the CRS and culture. These factors were, therefore, obviously not a source of heterogeneity among the studies.
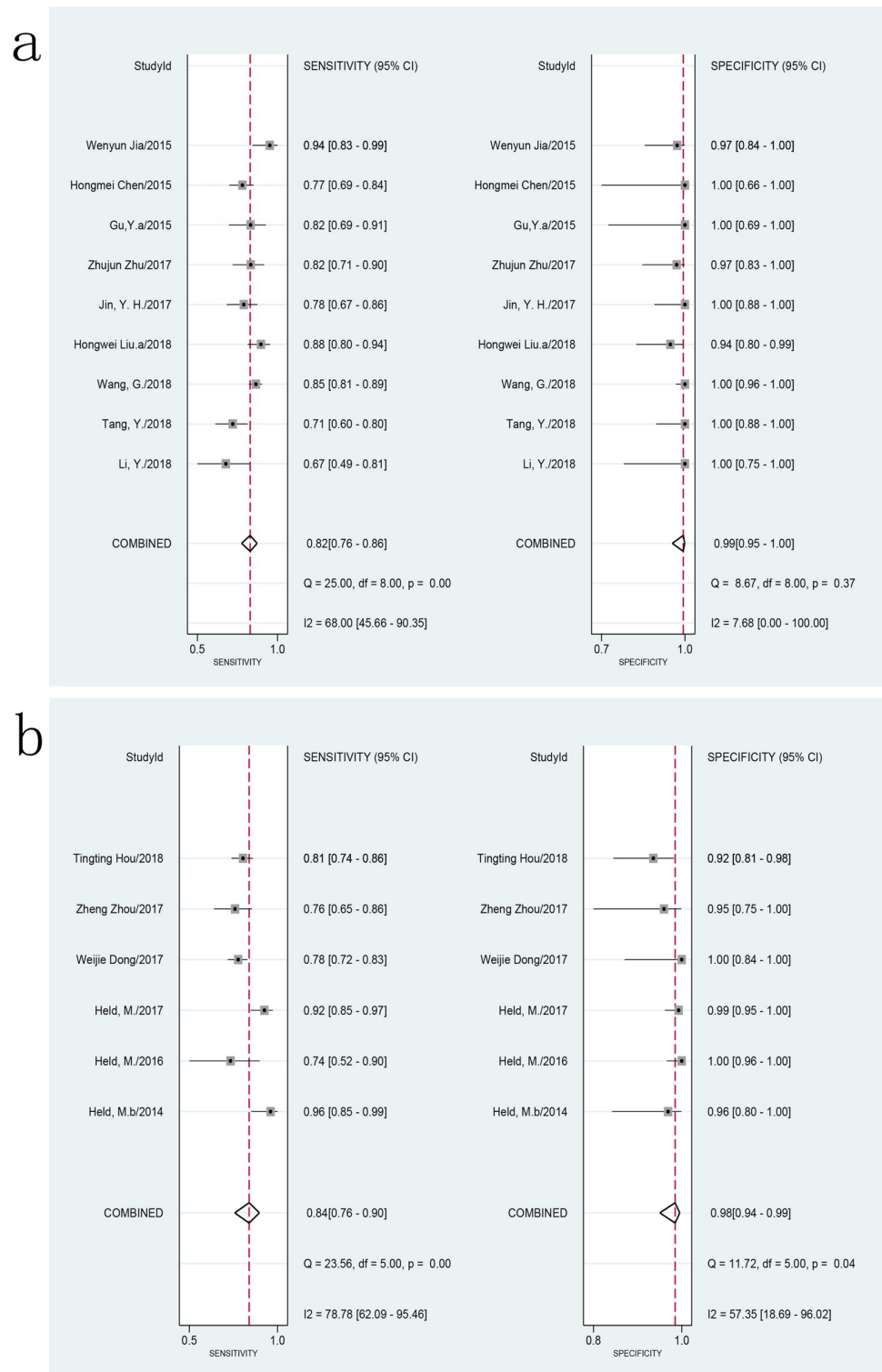
Fig 4. Forest plot of Xpert sensitivity and specificity for tuberculosis detection in BJTB compared with culture reference standard. BJTB: bone and joint tuberculosis.

https://doi.org/10.1371/journal.pone.0221427.g004

Discussion

The diagnosis of BJTB, just like that of other forms of EPTB, is very challenging, due to its paucibacillary nature [48]. The delayed diagnosis and treatment of BJTB can lead to serious complications, such as joint destruction, contractures in large joints, growth arrest, and neurological impairment in spinal tuberculosis, resulting in long-term morbidity and disability [36]. Invasive examination is a necessary diagnostic step in most cases. Puncture and biopsy of lesions are the most common invasive procedures for EPTB; the same applies to BJTB. Due to the deep location, obstruction of the bony structure, limited volume of the puncture specimen, the atypical nature of the puncture site, and low bacterial content of the specimen, it is still difficult to diagnose BJTB by pathological and cultural examination of the puncture specimen [39]. This causes BJTB to be often misdiagnosed or missed, leading to inappropriate treatment and adverse prognoses among patients. Therefore, in order to reduce the deformity and disability rate of BJTB, rapid and effective diagnostic methods are needed to detect BJTB early.

Nucleic acid tests such as the loop-mediated isothermal amplification assay, as a fast and efficient detection method, has been widely used in the diagnosis of tuberculosis [49] and demonstrated good diagnostic efficacy in the detection of EPTB [50]. The Xpert MTB/RIF assay, a rapid and automated real-time nucleic acid amplification test is currently one of the most commonly used in the diagnosis of tuberculosis. This test can validate the MTB complex DNA within 2 hours and is widely used in the diagnosis of PTB and EPTB. Previous studies have shown that this test has good diagnostic efficacy in the diagnosis of PTB and EPTB [7]. The Xpert MTB/RIF assay has also been recommended by the World Health Organization for the diagnosis of EPTB [51], including lymph node tuberculous and tuberculous meningitis. This

Fig 5. Forest plot of Xpert sensitivity and specificity for tuberculosis detection in BJTB vs. composite reference standard. a: pus samples. b: tissue samples. BJTB: Bone and joint tuberculosis.

https://doi.org/10.1371/journal.pone.0221427.g005
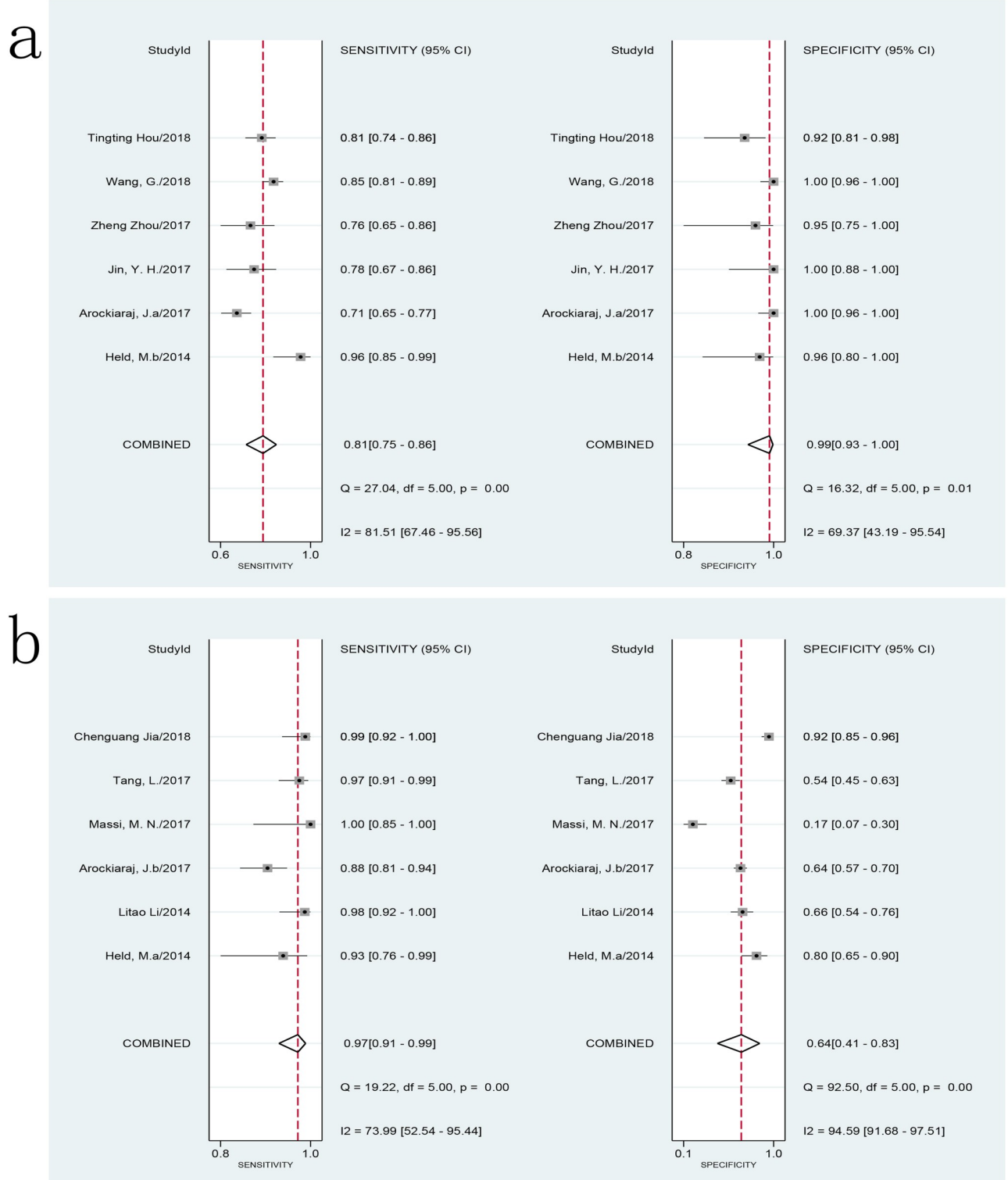
Fig 6. Forest plot of Xpert sensitivity and specificity for tuberculosis detection in BJTB for spinal tuberculosis. a: composite reference standard. b: culture reference standard. BJTB: Bone and joint tuberculosis.

https://doi.org/10.1371/journal.pone.0221427.g006

test should also be used in the diagnosis of BJTB. Although many studies had reported the application of this test in BJTB, there has been no consensus on the sensitivity and specificity. In a meta-analysis performed by Wen et al., the pooled sensitivity and specificity of Xpert MTB/RIF for BJTB were 81% (95% CI, 78–83) and 83% (95% CI, 80–86), respectively [52]. However, this result was not differentiated according to different reference standards; the number of studies included was relatively small, and the diagnostic efficacy of different types of specimens and different sites of infection was not differentiated. In order to further evaluate the efficacy of this test in the diagnosis of BJTB, we designed this study to determine the accuracy of Xpert MTB/RIF for BJTB.

Our study included 19 studies with comparisons to the CRS and 14 studies with comparisons to culture. The selection of patients included in most studies was not consecutive. The CRS varied among the articles included, which could be the main source of bias among the studies. Publication bias needed to be considered. For intervention studies, publication bias occurs if studies with significant results are more likely to be published than studies with non-significant findings. Regarding diagnostic tests, many studies are conducted without ethical review or study registration, or do not compare tests; hence, it would be problematic to assess publication bias. According to the Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy, traditional analytical approaches, such as 'funnel plots', may not be appropriate for the assessment of test accuracy [10]. At present, there is no recognized and accepted statistical method for quantifying the potential effect of publication bias in studies of diagnostic accuracy [53]. Therefore, we chose not to carry out a formal assessment of publication bias.

The present study demonstrated that the pooled sensitivity and specificity of Xpert MTB/RIF for BJTB were 81% (95% CI, 77–84) and 99% (95% CI, 97–100) compared to the CRS, 96% (95% CI, 90–98) and 85% (95% CI, 57–96) compared to culture, respectively. As expected, the use of different reference standards led to different sensitivities and specificities, and their ranges were relatively credible. Regardless of the gold standard, Xpert MTB/RIF showed a very good diagnostic accuracy. In the present study, a substantial level of heterogeneity was also observed among the studies. Subgroup analysis revealed that the pooled sensitivity of Xpert MTB/RIF performed on pus samples was 82% (I^2 = 68%; 95% CI, 76–86), and the pooled specificity was 99% (I^2 = 8%; 95% CI, 95–100) compared to the CRS, respectively. The pooled sensitivity and specificity were 84% (I^2 = 79%; 95% CI, 76–90) and 98% (I^2 = 57%; 95% CI, 94–99), respectively, compared to the CRS, when using tissue specimens. The level of heterogeneity within subgroups decreased, suggesting that specimen types might be a source of heterogeneity among studies especially for specificity. The sensitivity and specificity of Xpert MTB/RIF performed using pus specimens was similar to those obtained using tissue specimens compared to the CRS; there was no significant difference. Overall, the diagnostic accuracy of Xpert MTB/RIF for BJTB, using pus and tissue specimens, were found to be similar. This finding was similar to that made in a previous study of lymph node tuberculosis [8]. For spinal tuberculosis, the pooled sensitivity and specificity of Xpert MTB/RIF were 81% (95% CI, 75–86) vs. 97% (95% CI, 91–99) and 99% (95% CI, 93–100) vs. 64% (95% CI, 41–83) compared to CRS and culture, respectively. However, the heterogeneity (such as regards sample type) across studies was substantial; the results needed to be treated with caution. Owing to its paucibacillary nature, for BJTB, culture is an imperfect reference standard. A CRS with multiple evaluation indicators might be a more applicable reference standard. We chose to evaluate the two reference standards separately as it would be more confusing to mix them. As predicted, we found an underestimated level of sensitivity and an overestimated level of specificity when using the CRS as a reference standard. The range for sensitivity and specificity compared with the two references was more plausible. However, the CRS varied across studies in this study. The CRS for most studies included the results of culture (Lowenstein-Jensen and/or BACTEC MGIT

960 culture), except for four studies [20, 31, 32, 38] while that for most studies included the results of histology/cytology, smear microscopy, clinical symptoms, and radiographic features; eight studies included the response to anti-tuberculosis treatment, and none of the studies included other molecular test results. This might be one of the sources of heterogeneity among the studies. The studies also used different culture references. Nine studies only used BACTEC MGIT 960 liquid culture as the reference, one study only used Lowenstein-Jensen solid culture as the reference [34], and four studies used both of them as references [18,22,24,35]. The performances of BACTEC MGIT 960 liquid and Lowenstein-Jensen solid cultures were different [54], which might also be one of the sources of heterogeneity among the studies. Separate data used in the subgroup of different reference standards were limited; thus, further meta-analysis could not be performed.

Meta-regression analysis showed that the patient selection method affected the outcome and was a source of heterogeneity compared to the CRS, probably due to the fact that patients who were not included consecutively were more likely to introduce selection bias. The sample processing of BJTB specimens, such as decontamination, sample condition, and homogenization, varied among studies; however, meta-regression analysis showed that these factors did not affect the outcome and hence were not obvious sources of heterogeneity.

Our meta-analysis had several limitations. This meta-analysis has not been registered online and despite comprehensive searches, some studies may still have been missed, and some studies failed to distinguish specimen types. In addition, some studies used multiple sample types, which may have led to some bias in our results. Additionally, sample processing of BJTB specimens was highly variable among studies, since the assay, designed for respiratory samples, may slightly vary for other specimens. Additionally, the CRS standard for the studies was different. There was substantial heterogeneity among the studies, and the pooled estimates need to be interpreted with caution.

Conclusions

We observed that the pooled sensitivity and specificity of Xpert MTB/RIF were 81% and 99%, respectively, when compared to the CRS, and 96% and 85%, respectively, when compared to culture. For spinal tuberculosis, the pooled sensitivity and specificity of Xpert MTB/RIF were 81% and 99%, respectively, when compared with the CRS, and 97% and 64%, respectively, when compared with culture. When performed on pus samples, the pooled sensitivity and specificity were 82% and 99% compared with the CRS. When performed on tissue samples, the pooled sensitivity and specificity were 84% and 98% compared with the CRS. The diagnostic efficiencies for different specimen types (pus and tissue) were similar. Xpert MTB/RIF showed a good diagnostic accuracy for BJTB and was not related to the type of specimen.

Supporting information

S1 Supporting Information Data. analyzed in this study, including search strategy and funding. All data analyzed in this study are included in this published article and its supplementary information file.
(ZIP)

S1 File. PRISMA checklist.
(DOC)

Acknowledgments

Author Contributions

Conceptualization: Fangming Zhong, Xiaohua Kong.

Data curation: Yanqin Shen, Guocan Yu.

Formal analysis: Yanqin Shen, Xiaohua Kong.

Funding acquisition: Fangming Zhong.

Methodology: Yanqin Shen, Guocan Yu, Xiaohua Kong.

Resources: Fangming Zhong.

Software: Yanqin Shen, Guocan Yu.

Supervision: Fangming Zhong.

Writing – original draft: Yanqin Shen, Guocan Yu.

Writing – review & editing: Yanqin Shen, Guocan Yu, Xiaohua Kong.

References

1. World Health Organization. Global tuberculosis report 2018. Geneva, Switzerland; 2018.

2. Kulchavenya E. Extrapulmonary tuberculosis: are statistical reports accurate? Ther Adv Infect Dis. 2014; 2(2):61–70. https://doi.org/10.1177/2049936114528173 PMID: 25165556

3. Rajasekaran S, Soundararajan DCR, Shetty AP, Kanna RM. Spinal tuberculosis: Current concepts. Global spine J. 2018; 8(4 Suppl):96s–108s. https://doi.org/10.1177/2192568218769053 PMID: 30574444

4. Singh P, Saket VK, Kachhi R. Diagnosis of TB: From conventional to modern molecular protocols. Front Biosci (Elite Ed). 2019; 11:38–60.

5. Theron G, Peter J, Calligaro G, Meldau R, Hanrahan C, Khalfey H, et al. Determinants of PCR performance (Xpert MTB/RIF), including bacterial load and inhibition, for TB diagnosis using specimens from different body compartments. Sci Rep. 2014; 4:5658. https://doi.org/10.1038/srep05658 PMID: 25014250

6. Peirse M, Houston A. Extrapulmonary tuberculosis. Medicine. 2017; 45(12):747–752.

7. Brownell R, Metcalfe J, Millman AJ, Miller C, Cattamanchi A. Performance of Xpert MTB/RIF for diagnosis of pulmonary and extra-pulmonary tuberculosis-a systematic review and meta-analysis. Am J Respir Crit Care Med. 2012; 185:A4715.

8. Yu G, Zhong F, Ye B, Xu X, Chen D, Shen Y. Diagnostic accuracy of the Xpert MTB/RIF assay for lymph node tuberculosis: a systematic review and meta-analysis. Biomed Res Int. 2019; 2019:1–12.

9. 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. Ann Intern Med. 2011; 155(8):529–536. https://doi.org/10.7326/0003-4819-155-8-201110180-00009 PMID: 22007046

10. Macaskill P, Gatsonis C, Deeks JJ, Harbord RM, Takwoingi Y. Cochrane handbook for systematic reviews of diagnostic test accuracy version 1.0, The Cochrane Collaboration, 2010. http://srdta.cochrane.org/.

11. Zhou Y, Dendukuri N. Statistics for quantifying heterogeneity in univariate and bivariate meta-analyses of binary data: The case of meta-analyses of diagnostic accuracy. Stat Med. 2014; 33(16):2701–2717. https://doi.org/10.1002/sim.6115 PMID: 24903142

12. Schiller I, van Smeden M, Hadgu A, Libman M, Reitsma JB, Dendukuri N. Bias due to composite reference standards in diagnostic accuracy studies. Stat Med. 2016; 35(9):1454–1470. https://doi.org/10.1002/sim.6803 PMID: 26555849

13. Reitsma JB, Rutjes AW, Khan KS, Coomarasamy A, Bossuyt PM. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. J Clin Epidemiol. 2009; 62(8):797–806. https://doi.org/10.1016/j.jclinepi.2009.02.005 PMID: 19447581

14. Naaktgeboren CA, Bertens LCM, van Smeden M, de Groot JAH, Moons KGM, Reitsma JB. Value of composite reference standards in diagnostic research. BMJ. 2013; 347:f5605. https://doi.org/10.1136/bmj.f5605 PMID: 24162938

15. Held M, Laubscher M, Zar HJ, Dunn RN. GeneXpert polymerase chain reaction for spinal tuberculosis: an accurate and rapid diagnostic test. Bone Joint J. 2014; 96-B(10):1366–1369. https://doi.org/10.1302/0301-620X.96B10.34048 PMID: 25274923

16. Litao L, Hongmin L, Yuanzhen M, Dawei L, Da B, Xiaobo L, et al. Assessment of the Xpert MTB/RIF assay for the detection of M.tuberculosis and rifampicin resistance in spinal tuberculosis specimens.A retrospective validation study. Chinese Journal of Orthopaedics. 2014; 34(2):211–215. Chinese.

17. Gu Y, Wang G, Dong W, Li Y, Ma Y, Shang Y, et al. Xpert MTB/RIF and GenoType MTBDRplus assays for the rapid diagnosis of bone and joint tuberculosis. Int J Infect Dis. 2015; 36:27–30. https://doi.org/10.1016/j.ijid.2015.05.014 PMID: 26004172

18. Kim YW, Kwak N, Seong MW, Kim EC, Yoo CG, Kim YW, et al. Accuracy of the Xpert(R) MTB/RIF assay for the diagnosis of extra-pulmonary tuberculosis in South Korea. Int J Tuberc Lung Dis. 2015; 19(1):81–86. https://doi.org/10.5588/ijtld.14.0500 PMID: 25519795

19. Chen Hongmei, Li Xuelian, Gao Mengqiu, Fan Jun, Zhang Liqun. Assessment of the Xpert MTB/RIF assay for the detection of bone and joint tuberculous. Chinese Journal for Clinicians. 2015;(10):24–26. Chinese.

20. Jia Wenyun, Li Yuan. The application of Xpert MTB/RIF assay for rapid diagnosis of osteoarticular tuberculosis. Chinese Journal of Spine and Spinal Cord. 2015; 25(3):208–12. Chinese.

21. Held M, Laubscher M, Mears S, Dix-Peek S, Workman L, Zar H. Diagnostic accuracy of the Xpert MTB/RIF assay for extrapulmonary tuberculosis in children with musculoskeletal infections. Pediatr Infect Dis J. 2016; 35(11):1165–1168. https://doi.org/10.1097/INF.0000000000001271 PMID: 27286562

22. Zhen Li. Detection of Mycobacterium tuberculosis and rifampicin resistance in bone and joint tuberculosis lesions by Xpert MTB/RIF. Shanxi Medical University; 2016. Chinese.

23. Tang Kai, Dong Weijie, Lan Tinglong, Fan Jun, Li Yuan, Yan Guangxuan, et al. Study on rapid diagnosis of knee joint tuberculosis by Xpert MTB/RIF. Chinese Journal of Antituberculosis. 2016; 38(4):300–4. Chinese.

24. Arockiaraj J, Michael JS, Amritanand R, David KS, Krishnan V. The role of Xpert MTB/RIF assay in the diagnosis of tubercular spondylodiscitis. Eur Spine J. 2017; 26(12):3162–3169. https://doi.org/10.1007/s00586-017-5076-9 PMID: 28391384

25. Held M, Laubscher M, Workman L, Zar HJ, Dunn R. Diagnostic accuracy of GeneXpert MTB/RIF in musculoskeletal tuberculosis: High sensitivity in tissue samples of HIV-infected and HIV-uninfected patients. S Afr Med J. 2017; 107(10):854–8. https://doi.org/10.7196/SAMJ.2017.v107i10.11186 PMID: 29022528

26. Jin YH, Shi SY, Zheng Q, Shen J, Ying XZ, Wang YF. [Application value of Xpert MTB/RIF in diagnosis of spinal tuberculosis and detection of rifampin resistance]. Zhongguo Gu Shang. 2017; 30(9):787–791. https://doi.org/10.3969/j.issn.1003-0034.2017.09.002 PMID: 29455477

27. Li Y, Pang Y, Zhang T, Xian X, Wang X, Yang J, et al. Rapid diagnosis of extrapulmonary tuberculosis with Xpert Mycobacterium tuberculosis/rifampicin assay. J Med Microbiol. 2017; 66(7):910–914. https://doi.org/10.1099/jmm.0.000522 PMID: 28708053

28. Massi MN, Biatko KT, Handayani I, Pratama MY, Septriani S, Nurdin GM, et al. Evaluation of rapid GeneXpert MTB/RIF method using DNA tissue specimens of vertebral bones in patients with suspected spondylitis TB. J Orthop. 2017; 14(1):189–191. https://doi.org/10.1016/j.jor.2016.12.003 PMID: 28115796

29. Tang L, Feng S, Gao R, Han C, Sun X, Bao Y, et al. A comparative study on the role of Xpert MTB/RIF in testing different types of spinal tuberculosis tissue specimens. Genet Test Mol Biomarkers. 2017; 21(12):722–726. https://doi.org/10.1089/gtmb.2017.0149 PMID: 29172708

30. Dong Weijie, Qin Shibing, Lan Tinglong, Fan Jun, Tang Kai, Li Yuan, et al. Study on GeneXpert MTB/RIF for rapid diagnosis of bone and joint tuberculosis and detection of rifampin resistance. Chinese Journal of Antituberculosis. 2017; 39(4):337–341. Chinese.

31. Zhou Z, Wang X, Chen H, Ma Y, Zhou X, Jing H, et al. Application of GeneXpert MTB/RIF in the diagnosis of spinal tuberculosis. Shandong Medical Journal. 2017; 57(17):90–92. Chinese.

32. Zhu Z, Xia Q, Yue Y, Zhang Y, Hu J, Zhu M. Application value of Xpert MTB/RIF in diagnosis of osteoarthritis tuberculosis. Journal of Preventive Medicine. 2017; 29(3):322–324. Chinese.

33. Khan AS, Ali S, Khan MT, Ahmed S, Khattak Y, Abduljabbar, et al. Comparison of GeneXpert MTB/RIF assay and LED-FM microscopy for the diagnosis of extra pulmonary tuberculosis in Khyber Pakhtunkhwa, Pakistan. Braz J Microbiol. 2018; 49(4):909–913. https://doi.org/10.1016/j.bjm.2018.02.011 PMID: 29752010

34. Li Y, Jia W, Lei G, Zhao D, Wang G, Qin S. Diagnostic efficiency of Xpert MTB/RIF assay for osteoarticular tuberculosis in patients with inflammatory arthritis in China. PloS One. 2018; 13(6):e0198600. https://doi.org/10.1371/journal.pone.0198600 PMID: 29856840

35. Perez-Risco D, Rodriguez-Temporal D, Valledor-Sanchez I, Alcaide F. Evaluation of the Xpert MTB/RIF ultra assay for direct detection of Mycobacterium tuberculosis complex in smear-negative extrapulmonary samples. J Clin Microbiol. 2018; 56(9):e00659–18. https://doi.org/10.1128/JCM.00659-18 PMID: 29950333

36. Tang Y, Yin L, Tang S, Zhang H, Lan J. Application of molecular, microbiological, and immunological tests for the diagnosis of bone and joint tuberculosis. J Clin Lab Anal. 2018; 32(2). https://doi.org/10.1002/jcla.22260 PMID: 28543876

37. Wang G, Dong W, Lan T, Fan J, Tang K, Li Y, et al. Diagnostic accuracy evaluation of the conventional and molecular tests for spinal tuberculosis in a cohort, head-to-head study. Emerg Microbes Infect. 2018; 7(1):109. https://doi.org/10.1038/s41426-018-0114-1 PMID: 29921972

38. Hou T, Hou X, Li Y, Li D, Suo H, Zhou Z. Diagnostic value of GeneXpert MTB/RIF in detecting spinal tuberculosis. Laboratory Medicine and Clinic. 2018; 15(21):3268–3270. Chinese.

39. Jia C, Gao J, Yao X, Wang L, Yao L, Li W. The value of Xpert MTB/RIF in the early diagnosis of suspected spinal tuberculosis and the detection of rifampicin resistance. Journal of Hebei Medical University. 2018; 39(9):1031–1033. Chinese.

40. Liu H, Li X, Huang S, Liang G. Diagnostic Value of GeneXpert MTB/RIF detection in joint tuberculosis. Chinese Journal of Medicine. 2018; 53(6):688–91. Chinese.

41. Rao J, Cai Y, Cao G, Yang Z. Application of Xpert MTB/RIF in rapid diagnosis of spinal tuberculosis and detection of rifampin resistance. Chinese Journal of Trauma and Disability Medicine. 2019; 27(7):32–33. Chinese.

42. Kanade S, Nataraj G, Mehta P. Improved case detection using Xpert Mycobacterium tuberculosis/rifampicin assay in skeletal tuberculosis. Indian J Med Microbiol. 2018; 36(4):590–593. https://doi.org/10.4103/ijmm.IJMM_19_10 PMID: 30880713

43. Liu R, Gao M, Wu X, Zhang L, Song Y, Ma L. Diagnositic value of Xpert mycobacterium tuberculosis/rifampicin assay and γ-interferon release assay for joint tuberculosis. Chinese Journal of Medicine. 2016; 51(6):70–2. Chinese.

44. Sharma A, Chhabra HS, Mahajan R, Chabra T, Batra S. Magnetic resonance imaging and GeneXpert: a rapid and accurate diagnostic tool for the management of tuberculosis of the spine. Asian Spine J. 2016; 10(5):850–856. https://doi.org/10.4184/asj.2016.10.5.850 PMID: 27790312

45. Chen ZF, Lao HL, Li XH, Wang J, Chen Q, Wang ZX, et al. Experimental study of GeneXpert(®) system in the diagnosis of extra-pulmonary tuberculosis. Zhonghua Jie He He Hu Xi Za Zhi. 2016; 39(7):529–533. https://doi.org/10.3760/cma.j.issn.1001-0939.2016.07.008 PMID: 27430924

46. Garcia P, Balcells ME, Castillo C, Miranda C, Geoffroy E, Roman JC, et al. [Evaluation of Xpert(R) MTB/RIF technique for Mycobacterium tuberculosis complex detection in extra-respiratory specimens]. Rev Chilena Infectol. 2017; 34(4):333–339. https://doi.org/10.4067/s0716-10182017000400333 PMID: 29165509

47. Diallo AB, Kollo AI, Camara M, Lo S, Ossoga GW, Mbow M, et al. [Performance of GeneXpert MTB / RIF((R)) in the diagnosis of extrapulmonary tuberculosis in Dakar: 2010–2015]. Pan Afr Med J. 2016; 25:129. https://doi.org/10.11604/pamj.2016.25.129.10065 PMID: 28292091

48. Broderick C, Hopkins S, Mack DJF, Aston W, Pollock R, Skinner JA, et al. Delays in the diagnosis and treatment of bone and joint tuberculosis in the United Kingdom. Bone Joint J. 2018; 100-B(1):119–124. https://doi.org/10.1302/0301-620X.100B1.BJJ-2017-0357.R1 PMID: 29305460

49. Greco S, Girardi E, Navarra A, Saltini C. Current evidence on diagnostic accuracy of commercially based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis. Thorax. 2006; 61(9):783–790. https://doi.org/10.1136/thx.2005.054908 PMID: 16738037

50. Yu G, Shen Y, Zhong F, Ye B, Yang J, Chen G. Diagnostic accuracy of the loop-mediated isothermal amplification assay for extrapulmonary tuberculosis: A meta-analysis. PLoS One. 2018; 13(6): e0199290. https://doi.org/10.1371/journal.pone.0199290 PMID: 29944682

51. Denkinger CM, Schumacher SG, Boehme CC, Dendukuri N, Pai M, Steingart KR. Xpert MTB/RIF assay for the diagnosis of extrapulmonary tuberculosis: a systematic review and meta-analysis. Eur Respir J. 2014; 44(2):435–446. https://doi.org/10.1183/09031936.00007814 PMID: 24696113

52. Wen H, Li P, Ma H, Lv G. Diagnostic accuracy of Xpert MTB/RIF assay for musculoskeletal tuberculosis: A meta-analysis. Infect Drug Resist. 2017; 10:299–305. https://doi.org/10.2147/IDR.S145843 PMID: 29026323

53. Wilson C, Kerr D, Noel-Storr A, Quinn TJ. Associations with publication and assessing publication bias in dementia diagnostic test accuracy studies. Int J Geriatr Psychiatry. 2015; 30(12):1250–1256. https://doi.org/10.1002/gps.4283 PMID: 25779466

54. Diriba G, Kebede A, Yaregal Z, Getahun M, Tadesse M, Meaza A, et al. Performance of Mycobacterium Growth Indicator Tube BACTEC 960 with Lowenstein-Jensen method for diagnosis of Mycobacterium tuberculosis at Ethiopian National Tuberculosis Reference Laboratory, Addis Ababa, Ethiopia. BMC Res Notes. 2017; 10(1):181. https://doi.org/10.1186/s13104-017-2497-9 PMID: 28486950