

# Pooled analysis of the Xpert MTB/RIF assay for diagnosing tuberculous meningitis

Yuan-Zhi Chen <sup>1#</sup>, Li-Chang Sun <sup>1#</sup>, Yao-Hong Wen <sup>1#</sup>, Zhong-Wei Li <sup>1</sup>, Shu-Jin Fan <sup>1</sup>, Hong-Kun Tan <sup>1</sup>, Min Qiu <sup>1</sup>, Zhi-Yong Pan <sup>1</sup>, Qin Li <sup>5</sup>, Yan-Zhen Zhao <sup>2</sup>, Zhen-Xing Li <sup>1</sup>, Xu-Guang Guo <sup>1, 2, 3, 4\*</sup>

1 Department of Clinical Medicine, the Third Clinical School of Guangzhou Medical University, Guangzhou 511436, China; 2 Department of Clinical Laboratory Medicine, the Third affiliated hospital of Guangzhou Medical University, Guangzhou 510150, China; 3 Key Laboratory for Major Obstetric Diseases of Guangdong Province, Guangzhou 510150, China; 4 Key Laboratory of Reproduction and Genetics of Guangdong Higher Education Institutes, Guangzhou 510150, China; 5 Section of Pulmonary, Critical Care and Sleep Medicine, Department of Medicine, Yale School of Medicine, Yale University, New Haven, CT 06520-8057, USA

<sup>#</sup>Contributed equally to this work.

\*Corresponding author:

Xu-Guang Guo: [gysyngx@gmail.com](mailto:gysyngx@gmail.com)

## Abstract

**Background:** Tuberculous meningitis (TBM) is one of the most serious types of extrapulmonary tuberculosis. However, low sensitivity of culture of cerebrospinal fluid (CSF) increases the difficulty in clinical diagnosis, leading to diagnostic delay, and misdiagnosis. Xpert MTB/RIF assay is a rapid and simple method to detect tuberculosis. However, the efficacy of this technique in diagnosing TBM remains unclear. Therefore, a meta-analysis was conducted to evaluate the diagnostic efficacy of Xpert MTB/RIF for TBM, which may enhance the development of early diagnosis of tuberculous meningitis.

**Methods:** Relevant studies in the PubMed, Embase, and Web of Science databases were retrieved using the key words “Xpert MTB/RIF,” “tuberculous meningitis (TBM)”. The pooled sensitivity, pooled specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, summary receiver operator characteristic curve, and area under the curve (AUC) of Xpert MTB/RIF were determined and

analyzed.

**Results:** A total of 162 studies were enrolled and only 14 met the criteria for meta-analysis. The overall pooled sensitivity of Xpert MTB/RIF was 63% [95% confidence interval (CI), 59%–66%], while the overall pooled specificity was 98.1% (95% CI, 97.5%-98.5%). The pooled values of positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio were 20.91% (12.71%–52.82%), 0.40% (0.32%–0.50%), and 71.49% (32.64%–156.56%), respectively. The AUC was 0.76.

**Conclusions:** Xpert MTB/RIF exhibited high specificity in diagnosing TBM in CSF samples, but its sensitivity was relatively low. It is necessary to combine other high-sensitive detection methods for the early diagnosis of TBM. Moreover, the centrifugation of CSF samples was found to be beneficial in improving the sensitivity.

**Key words:** Diagnosis; meta-analysis; tuberculous meningitis; Xpert MTB/RIF

## Introduction

The World Health Organization (WHO) has reported tuberculosis as one of the top 10 leading causes of death worldwide<sup>1</sup>. It is an infectious disease caused by *Mycobacterium tuberculosis* (MTB) which causes pulmonary tuberculosis and extrapulmonary tuberculosis (EPTB). More than 50% of infected people either die or become disabled<sup>2</sup>. TBM is a kind of central nervous system tuberculosis caused by *M. tuberculosis* invading the brain through blood circulation or other means. It belongs to a paucibacillary disease group. In the conventional culture methods, the slow growth of *M. tuberculosis* and the low sensitivity of microscopic analysis of CSF hinder the laboratory diagnosis of TBM and increase the difficulty in clinical diagnosis, leading to diagnostic delay, misdiagnosis, and increased mortality<sup>3,4</sup>. Therefore, the early diagnosis of TBM is of important clinical significance in reducing its harmfulness<sup>5</sup>.

*MTB* culture is the traditional reference standard for diagnosing TBM, but it is not perfect. On the one hand, its sensitivity is moderate<sup>6</sup>; on the other hand, it requires strict technical operation, pollution-free environment, and long culture time. In 2010, Marais et al. proposed a composite reference standard (CRS)<sup>7</sup>, which combined clinical assessment and laboratory examination results. However, it easily led to false-positive results due to various factors, such as the subjectivity of clinical assessment. As a fast and simple detection method, the Xpert MTB/RIF assay not only has simple operation but also has low requirement for experimental environment and can obtain results within 2 h. As Xpert MTB/RIF is tested with a sealed disposable reaction tube, the risk of contamination and the resulting false-positive results can be reduced to almost negligible levels. However, a great controversy still exists about the diagnostic efficacy of Xpert MTB/RIF in TBM. Although several studies were conducted on the diagnostic performance of Xpert MTB/RIF test for TBM<sup>5,8-20</sup>, the results were contradictory. Hence, this meta-analysis was performed to evaluate the diagnostic efficacy of Xpert MTB/RIF for TBM.

## Methods

### Literature retrieval

The studies published before January 30, 2019, in PubMed, Embase, and Web of Science (WOS) were searched using the following key words: “Xpert MTB/RIF,” “tuberculous meningitis (TBM)”. The detailed search strategy see “annex 1”. Some relevant references were also retrieved. Language restriction was set to "English only".

The inclusion criteria were as follows: (1) culture or CRS(only meeting criteria on definite tuberculous meningitis (definite TBM): microbiological identification or evidence from commercial nucleic acid amplification tests of CNS *Mycobacterium tuberculosis* infection.) data were compared with Xpert MTB/RIF to determine the accuracy of diagnosing TBM; (2) human samples were analyzed; (3) enough data was generated to construct a  $2 \times 2$  table for calculating sensitivity, specificity, and likelihood ratio, and in studies with unreported data, the authors were requested for

the required data; and (4) the sample size of studies was not less than 15.

The exclusion criteria were as follows: (1) duplicate studies; (2) animal studies; (3) abstracts, conference abstracts, comments, reviews, letters, and case reports; and (4) lack of complete raw data, raw data not enough to construct  $2 \times 2$  tables, or unable to contact authors to obtain raw data.

### **Literature screening**

All the studies were screened and retrieved by the two researchers independently. After completing the independent screening, two researchers checked the screening results and discussed the inclusion or exclusion of inconsistent results through consultation. If the negotiation failed, a third researcher was assigned to conduct a screening evaluation; all the results were pooled together.

### **Data extraction**

Two researchers independently extracted data from the enrolled studies. These data included the following: (1) the basic information (such as author, year, and region); (2) reference standard; (3) CSF centrifugation operation; (4) data included in the fourfold tables of the study, such as the total number of study samples, true positive, false positive (FP), false negative, and true negative; and (5) TBM prevalence (i.e., reference standard positive rate).

### **Quality assessment of the studies**

Two researchers independently evaluated the included studies using Quality assessment of diagnostic accuracy studies-2 (QUADAS-2) as a criterion<sup>21</sup>, including Risk of bias and applicability concern. Risk of bias contains 11 signal problems, and the risk in the four domains is determined by the answers to their corresponding signal questions.

Applicability concerns include three domains. In the domain of patient's selection, We scored “low concern” if patients were evaluated at local hospitals or primary

care centres. We scored “high concern” if patients were evaluated exclusively as inpatients at tertiary care centres. We scored “unclear concern” if the clinical setting was not reported or if information was insufficient to allow a decision.

In the domain of index test, We judged “low concern” if the index test was performed as recommended by the manufacturer for sputum. We scored “high concern” if the test was performed in a way that deviated from these recommendations. We scored “unclear concern” if we could not tell.

In the domain of reference standard We judged “high concern” if included studies did not speciate mycobacteria isolated in culture or the patient is not a definite TBM in the CRS, “low concern” if speciation was performed or the patient is a definite TBM in the CRS and “unclear concern” if we could not tell.

The Review Manager 5.2 software was used to display the quality of studies and make charts. Any disagreement arising from this process was communicated to a third researcher, and a consensus was reached.

### **Data analysis**

The Meta-Disc software v.1.4<sup>22</sup> was used to analyze the extracted fourfold table data: sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, diagnostic odds ratio, and its corresponding 95% confidence interval (CI). The accuracy of TBM diagnosis by Xpert MTB/RIF was analyzed by the stochastic effect model and presented in the form of a forest map. Deeks’ funnel plot was drawn using Stata software. Deeks’ funnel plot is mainly used to observe whether there are biases in the results of a systematic evaluation or meta-analysis, such as publication bias or other biases. If the test results were  $P < 0.01$ , the publication bias test results were significant.

### **Heterogeneity analysis**

Heterogeneity may arise from contingent opportunistic factors or from the threshold effect of defining negative and positive results. The heterogeneity caused by opportunistic factors was explored through the following operations: (1) Visual

inspection of forest maps to observe the deviation between the sensitivity and specificity of each entry and the vertical lines of the corresponding pooled values; a large deviation between the study and the line indicated a possible source of heterogeneity. (2) The chi-square *P* value generated automatically by the Meta-Disc software was used to judge the data. A lower chi-square *P* value indicated that heterogeneity originated from many sources and was not clear. (3) Quantitative indicators of heterogeneity were judged by the inconsistency index (*I*-square) automatically generated by the Meta-Disc software. The inconsistency index was interpreted as follows: 0%–40%: low heterogeneity; 50%–70%: moderate heterogeneity; and >70%: significant heterogeneity<sup>23</sup>.

The heterogeneity caused by the threshold effect was explored by drawing summary receiver operating characteristic (SROC) curves of Xpert MTB/RIF to evaluate whether the points in the curve had a curved (shoulder-arm) pattern. Typical “shoulder-arm” patterns indicated the presence of a threshold effect<sup>22</sup>. The Meta-Disc software automatically calculated and displayed the area under SROC curve and the Cochrane index (*Q*\*). Cochrane index, as well as SROC curve, which is a comprehensive index reflecting continuous variables of sensitivity and specificity, and reveals the relationship between sensitivity and specificity by composition. As a further evaluation of the threshold effect, the Spearman correlation coefficient was also calculated. Spearman correlation coefficient is a nonparametric measure of the dependence of two variables. It uses monotone equation to evaluate the correlation between two statistical variables. If the Spearman correlation coefficient was more than 0.6, the possibility of threshold effect was indicated. If the value was less than 0.6, the absence of threshold effect was indicated.

## Results

### Literature screening and inclusion process

A total of 162 studies (44 in PubMed, 45 in Embase, and 73 in WOS) were searched. Of these, 57 duplicate studies were excluded. Of the remaining 105 studies, 80 studies were excluded (11 case reports, 18 reviews, 13 letters, 8 meeting abstracts, 7 comments, and 23 unrelated studies) by browsing the titles and abstracts of the studies according to the inclusion/exclusion criteria. By browsing the full text of the remaining 25 studies, 11 studies were further excluded because of the following

reasons: 2 studies lacked the reference standard such as culture or CRS, the sample sizes of 2 studies were less than 15, and the original data of 7 studies could not form a complete  $2 \times 2$  table. 0 grey literature was found during the second screening and the full-text browsing. Finally, 14 studies that met the inclusion/exclusion criteria were included (Fig. 1).

### **Characteristics and quality assessment of included studies**

A total of 14 studies and 4147 samples were included, and 20 complete  $2 \times 2$  tables were extracted from them. We extracted 20 sets of data from these 14 studies, of which 7 using CRS as a reference standard and 13 using culture as a reference standard. In addition, we divided these studies into cerebrospinal fluid centrifugation group (15 sets of data), non-centrifugation group (4 sets of data) and mixed group (1 set of data, which contains the data of both centrifuged and non-centrifuged CSF sample.). The features included in the study are summarized in “Table 1”.

“Figures 2” and 3 show the quality assessment of 14 included studies, including bias risk and applicability concerns. In the patient selection domain, 13 studies (92.85%) had a lower risk of bias and 1 study (7.15%) was rated as having “unclear” risk of bias because patient selection was not clear. The applicability concerns of eight studies (57.14%) were rated as “high concern” because their patients were evaluated in tertiary hospital centers rather than in local community hospitals or in primary hospitals. Six studies (42.86%) were rated as “unknown concerns” because it was not clear whether their patients were assessed in local primary hospitals or tertiary hospitals. In the index test domain, all the bias risks included in the studies were rated as “low risk”; and the applicability concerns of two studies (14.25%) were rated as “unknown concern” because the details of sample processing were not reported, and the rest were “low concern”. In the reference standard domain, all the bias risks included in the studies were rated “low risk” and their applicability concerns were “low concern.” In the flow and timing domain, all the studies were judged to be a low-bias risk because all patients were included in the analysis using appropriate reference criteria and with appropriate intervals between index tests and reference criteria.

### **Publication bias**

No publication bias was found in the Deeks' funnel plot (Fig. 4). The experimental results  $P = 0.029$  ( $>0.01$ ) indicated no publication bias.

### **Spearman correlation coefficient for threshold effect analysis**

In the threshold effect analysis, the Spearman correlation coefficient was 0.014 ( $<0.6$ ) and  $P$  value was 0.952 ( $>0.05$ ). In addition, the SROC curve was analyzed (Fig. 5), which had no “shoulder-arm” distribution characteristics. Therefore, it was concluded that no threshold effect existed in the included studies.

### **Summary receiver operator characteristic curve**

The area under the SROC curve (Fig. 5) and the Cochrane ( $Q^*$ ) index was 0.76 and 0.70, respectively, indicating that TBM detection by Xpert MTB/RIF had relatively moderate accuracy.

### **Meta-analysis results**

The sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, and diagnostic odds ratio (Figs. 6–10) and their 95% CI were 63% (59%–66%), 98.1% (97.5%–98.5%), 20.91% (12.71%–52.82%), 0.40% (0.32%–0.50%), and 71.49% (32.64%–156.56%), respectively. The results of the pooled values suggested that the accuracy of TBM detection by Xpert MTB/RIF was relatively moderate.

Overall, visual forest maps showed a large deviation between the sensitivity of more than half of the included documents and the vertical line of the pooled values, indicating a greater possibility of heterogeneity. In terms of specificity, only one document deviated from the pooled value greatly, presumably due to some unknown contingency factors, while the rest were mostly concentrated near the pooled value. The chi-square  $P$  values of all the pooled values were very low, and the  $I$ -square values were higher than 70%.

### **Subgroup analysis**

Subgroup analysis was performed (Table 2) according to whether the standard reference was different (group A); whether CSF samples were centrifuged (group B); and whether the prevalence of TBM was more than 30% (i.e., positive rate; group C) Group A: When culture was used as the reference standard, the pooled sensitivity was 67% (63%–71%) and the pooled specificity was 98% (97%–98%). When CRS was used as the reference standard, the pooled sensitivity was 56% (51%–61%) and the pooled specificity was 99% (99%–100%).

Group B: When CSF samples were centrifuged, the pooled sensitivity was 63% (60%–67%) and the pooled specificity was 99% (98%–99%). When CSF samples were not centrifuged, the pooled sensitivity was 57% (48%–65%) and the pooled specificity was 95% (92%–97%).

Group C: When the prevalence of TBM was more than 30%, the pooled sensitivity was 62% (57%–66%) and the pooled specificity was 95% (93%–97%). When the prevalence of TBM was less than 30%, the pooled sensitivity was 64% (59%–69%) and the pooled specificity was 99% (98%–99%).

## Discussion

This study showed that the pooled specificity of Xpert MTB/RIF in diagnosing TBM was particularly high (98%), indicating that the possibility of misdiagnosis of TBM was low. The pooled sensitivity was 63% (59%–66%), which was not ideal and unlikely to greatly improve the accuracy of early diagnosis of TBM.

Comparing the data of the two groups with culture and CRS as the reference standards, it was found that although the pooled specificity of the two groups was not significantly different (98% and 99%, respectively, –1%), the pooled sensitivity of the two groups was quite different (67% and 56%, 11%). The decrease in sensitivity indicated that culture might have FP results, increasing the positive results, or that culture cannot be used as a single reference standard. This coincided with the conclusions of the WHO Steering Group (Christopher Gilpin et al.)<sup>24</sup>. When it comes to the discussion on Publication bias, there are two arguments about the defined value

of the P value, one is 0.01 and the other is 0.05, depending on the accuracy of the decision. We made Deeks' funnel plot by using Stata software. It is a default setting in the stata software: When the P value is greater than 0.01, there is no publication bias. The P value of our funnel plot is 0.029 ( $>0.01$ ), indicating no publication bias. However, it is generally believed that when  $P < 0.01$ , publication bias is very significant, and when  $P > 0.05$ , it indicates that publication bias does not exist. When P is between 0.01 and 0.05, it can be considered that there is a certain publication bias. So we did a sensitivity analysis by manually deleting the documents one by one. With each of the studies individually removed, the corresponding pooled values and the I-square were not materially altered for all models. It indicated that no single study influenced the pooled value and the I-square qualitatively, suggesting that the results of this meta-analysis are stable.

With a culture as the reference standard, the sensitivity was 12.5% lower than that indicated by Christopher Gilpin et al. (79.5%, 62.0%–90.2%), although the specificity of this study was similar (98.6%, 95.8%–99.6%). The reasons might be as follows: (1) compared with the data before 2013, the present study included more studies published in the last 6 years. (2) The inclusion and exclusion criteria defined in this study had two requirements: the sample size should be more than 15, and the fourfold table should be complete (i.e., only when at most one of the four data, true positive, false positive, true negative, false negative, is 0, that makes sense). Christopher Gilpin et al. included a lot of studies that did not meet the aforementioned conditions, leading to the inclusion of part of the studies and hence a smaller number of samples, besides the lack of sensitivity and specificity. It was believed that such an operation would increase the risk of experimental bias, which was not explained clearly by Christopher Gilpin and others.

When using CRS as the reference standard, the pooled sensitivity and specificity of this study were not significantly different from those of Christopher Gilpin and others (0.5% and 4%, respectively). That is, CRS as the reference standard might lead to an increase in the number of positive people, thus reducing the sensitivity of Xpert MTB/RIF.

In 2018, Kohli et al.<sup>25</sup> reviewed the diagnostic efficacy of Xpert MTB/RIF for EPTB.

They mentioned that the pooled sensitivity and 95% CI of Xpert MTB/RIF for TBM (taking culture as the reference standard) were 71.1% and 60.9%–80.4%, and the pooled specificity was 98.0% (97.0%–98.8%). The pooled specificity of the two methods was the same, and the pooled sensitivity was 4.1%. The reasons for this were the same as the aforementioned differences from the study by Christopher Gilpin et al.

When CSF specimens were centrifuged, both sensitivity and specificity were improved by 6% and 4%, respectively. This suggested that the centrifugation of CSF samples could improve the sensitivity of Xpert MTB/RIF, which might be due to the concentration of *M. tuberculosis* by centrifugation.

The difference in the pooled susceptibility and specificity of Xpert MTB/RIF between low-TBM epidemic areas and high-TBM epidemic areas was not significant, indicating that the prevalence of TBM might not be the source of heterogeneity.

Our study has the following limitations: first, clinical data were still relatively small, and the impact of unpublished positive results on the overall results could not be ruled out. In addition, few studies included the age and human immunodeficiency virus status of patients, and it was uncertain what impact it would have on the pooled values. Moreover, the impact of other potential confounding factors on the results was not explored. Finally, things in studies such as “patient selection” and “whether the hospitals were primary or tertiary” were unclear without contacting the authors, so the characteristics and quality were not well illustrated.

In conclusion, Xpert MTB/RIF exhibited high specificity in diagnosing TBM in CSF samples; however, its sensitivity was relatively low. It is necessary to combine other sensitive detection methods for the early diagnosis of TBM. Moreover, the centrifugation of CSF samples was found to be beneficial in improving sensitivity. However, more clinical data are needed to support the conclusions.

## Acknowledgements

We express our sincere gratitude to all members of our study team for their whole-hearted cooperation and to the original authors of the studies included for their wonderful work and valuable participation.

## **Funding**

This work was supported by College students' scientific research project of the third affiliated hospital of Guangzhou medical university (No. 2018A0021).

## **Availability of data and materials**

All data generated or analyzed during this study are included in this published article and its supplementary information files.

## **Authors' contributions**

Xu-Guang Guo conceived and designed the experiments. Yuan-Zhi Chen, Li-Chang Sun, Yao-Hong wen and Zhong-Wei Li retrieved and reviewed the study reports, summarized and analysed the data, and prepared the manuscript. Zhi-Yong Pan, Qin Li, Yan-Zhen Zhao, Zhen-Xing Li, Shu-Jin Fan, Hong-Kun Tan and Min Qiu retrieved some of the study reports and critically revised the manuscript versions. All authors read and approved the final manuscript.

## **Competing interests**

The authors declare that they have no competing interests.

## **Consent for publication**

Not applicable.

## **Ethics approval and consent to participate**

Not applicable.

## **References**

- 1 World Health Organization Global tuberculosis report 2018. Available at: [http://www.who.int/tb/publications/global\\_report/en](http://www.who.int/tb/publications/global_report/en).
- 2 Thwaites, G. E. Advances in the diagnosis and treatment of tuberculous meningitis. *Current opinion in neurology* **26**, 295-300, doi:10.1097/WCO.0b013e3283602814 (2013).
- 3 Purohit, M. & Mustafa, T. Laboratory Diagnosis of Extra-pulmonary Tuberculosis (EPTB) in Resource-constrained Setting: State of the Art, Challenges and the Need. *Journal of clinical*

- and diagnostic research : *JCDR* **9**, Ee01-06, doi:10.7860/jcdr/2015/12422.5792 (2015).
- 4 Mandal, N., Anand, P. K., Gautam, S., Das, S. & Hussain, T. Diagnosis and treatment of paediatric tuberculosis: An insight review. *Crit Rev Microbiol* **43**, 466-480, doi:10.1080/1040841x.2016.1262813 (2017).
- 5 Nhu, N. T., Heemskerk, D., Thu do, D. A., Chau, T. T., Mai, N. T., Nghia, H. D. *et al.* Evaluation of GeneXpert MTB/RIF for diagnosis of tuberculous meningitis. *Journal of clinical microbiology* **52**, 226-233, doi:10.1128/jcm.01834-13 (2014).
- 6 Huaroto, L., Kasper, M., Ramos, E. S., Mori, N., Jittamala, P., Chotivanich, K. *et al.* Uniform research case definition criteria differentiate tuberculous and bacterial meningitis in children. *PLoS one* **59**, 1574-1578, doi:10.1371/journal.pone.0198695 10.1093/cid/ciu665 (2014).
- 7 Marais, S., Thwaites, G., Schoeman, J. F., Torok, M. E., Misra, U. K., Prasad, K. *et al.* Tuberculous meningitis: a uniform case definition for use in clinical research. *The Lancet. Infectious diseases* **10**, 803-812, doi:10.1016/s1473-3099(10)70138-9 (2010).
- 8 Bahr, N. C., Nuwagira, E., Evans, E. E., Cresswell, F. V., Bystrom, P. V., Byamukama, A. *et al.* Diagnostic accuracy of Xpert MTB/RIF Ultra for tuberculous meningitis in HIV-infected adults: a prospective cohort study. *The Lancet. Infectious diseases* **18**, 68-75, doi:10.1016/s1473-3099(17)30474-7 (2018).
- 9 Bahr, N. C., Tugume, L., Rajasingham, R., Kiggundu, R., Williams, D. A., Morawski, B. *et al.* Improved diagnostic sensitivity for tuberculous meningitis with Xpert((R)) MTB/RIF of centrifuged CSF. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* **19**, 1209-1215, doi:10.5588/ijtld.15.0253 (2015).
- 10 Cresswell, F. V., Bangdiwala, A. S. & Bahr, N. C. Can improved diagnostics reduce mortality from Tuberculous meningitis? Findings from a 6.5-year cohort in Uganda. **3**, 64, doi:10.12688/wellcomeopenres.14610.2 (2018).
- 11 Heemskerk, A. D., Donovan, J., Thu, D. D. A., Marais, S., Chaidir, L., Dung, V. T. M. *et al.* Improving the microbiological diagnosis of tuberculous meningitis: A prospective, international, multicentre comparison of conventional and modified Ziehl-Neelsen stain, GeneXpert, and culture of cerebrospinal fluid. *The Journal of infection* **77**, 509-515, doi:10.1016/j.jinf.2018.09.003 (2018).
- 12 Metcalf, T., Soria, J., Montano, S. M., Ticona, E., Evans, C. A., Huaroto, L. *et al.* Evaluation of the GeneXpert MTB/RIF in patients with presumptive tuberculous meningitis. *PLoS one* **13**, doi:10.1371/journal.pone.0198695 (2018).
- 13 Patel, V. B., Connolly, C., Singh, R., Lenders, L., Matinyenya, B., Theron, G. *et al.* Comparison of amplicor and GeneXpert MTB/RIF tests for diagnosis of tuberculous meningitis. *Journal of clinical microbiology* **52**, 3777-3780, doi:10.1128/jcm.01235-14 (2014).
- 14 Patel, V. B., Theron, G., Lenders, L., Matinyena, B., Connolly, C., Singh, R. *et al.* Diagnostic accuracy of quantitative PCR (Xpert MTB/RIF) for tuberculous meningitis in a high burden setting: a prospective study. *PLoS medicine* **10**, e1001536, doi:10.1371/journal.pmed.1001536 (2013).
- 15 Pink, F., Brown, T. J., Kranzer, K. & Drobniowski, F. Evaluation of Xpert MTB/RIF for Detection of Mycobacterium tuberculosis in Cerebrospinal Fluid. *Journal of clinical microbiology* **54**, 809-811, doi:10.1128/jcm.02806-15 (2016).

- 16 Rufai, S. B., Singh, A., Singh, J., Kumar, P., Sankar, M. M. & Singh, S. Diagnostic usefulness of Xpert MTB/RIF assay for detection of tuberculous meningitis using cerebrospinal fluid. *The Journal of infection* **75**, 125-131, doi:10.1016/j.jinf.2017.04.010 (2017).
- 17 Sharma, K., Sharma, M., Chaudhary, L., Modi, M., Goyal, M., Sharma, N. *et al.* Comparative evaluation of Xpert MTB/RIF assay with multiplex polymerase chain reaction for the diagnosis of tuberculous meningitis. *Tuberculosis (Edinburgh, Scotland)* **113**, 38-42, doi:10.1016/j.tube.2018.09.002 (2018).
- 18 Solomons, R. S., Visser, D. H., Friedrich, S. O., Diacon, A. H., Hoek, K. G., Marais, B. J. *et al.* Improved diagnosis of childhood tuberculous meningitis using more than one nucleic acid amplification test. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* **19**, 74-80, doi:10.5588/ijtld.14.0394 (2015).
- 19 Solomons, R. S., Visser, D. H., Marais, B. J., Schoeman, J. F. & van Furth, A. M. Diagnostic accuracy of a uniform research case definition for TBM in children: a prospective study. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease* **20**, 903-908, doi:10.5588/ijtld.15.0509 (2016).
- 20 Wang, T.,Feng, G., D. Pang, Y., Yang, Y. ,N., Dai, W. ,Zhang, L. *et al.* Sub-optimal Specificity of Modified Ziehl-Neelsen Staining for Quick Identification of Tuberculous Meningitis. *Frontiers in microbiology* **7**, 2096, doi:10.3389/fmicb.2016.02096 (2016).
- 21 Whiting, P. F., Rutjes, A. W., Westwood, M. E., Mallett, S., Deeks, J. J., Reitsma, J. B. *et al.* QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine* **155**, 529-536, doi:10.7326/0003-4819-155-8-201110180-00009 (2011).
- 22 Zamora, J., Abraira, V., Muriel, A., Khan, K. & Coomarasamy, A. Meta-DiSc: a software for meta-analysis of test accuracy data. *BMC medical research methodology* **6**, 31, doi:10.1186/1471-2288-6-31 (2006).
- 23 Identifying and measuring heterogeneity. Available at [http://handbook.cochrane.org/chapter\\_9/9\\_5\\_2\\_identifying\\_and\\_measuring\\_heterogeneity.htm](http://handbook.cochrane.org/chapter_9/9_5_2_identifying_and_measuring_heterogeneity.htm) .
- 24 in *Automated Real-Time Nucleic Acid Amplification Technology for Rapid and Simultaneous Detection of Tuberculosis and Rifampicin Resistance: Xpert MTB/RIF Assay for the Diagnosis of Pulmonary and Extrapulmonary TB in Adults and Children: Policy Update* (World Health Organization Copyright (c) World Health Organization 2013., 2013).
- 25 Kohli, M., Schiller, I., Dendukuri, N., Dheda, K., Denkinger, C. M., Schumacher, S. G. *et al.* Xpert((R)) MTB/RIF assay for extrapulmonary tuberculosis and rifampicin resistance. *The Cochrane database of systematic reviews* **8**, Cd012768, doi:10.1002/14651858.CD012768.pub2 (2018).

## Figure and Table legends

Figure 1. Flow diagram of study identification and inclusion.

Figure 2. Quality evaluation of the included studies.

Figure 3. Risk of bias and applicability concerns graph: review authors' judgments about each domain presented as percentages across the included studies.

Figure 4. Deeks' funnel plot asymmetry test to assess publication bias in estimates of diagnostic odds ratio for Xpert MTB/RIF detection of tuberculous meningitis.

Figure 5. Summary receiver operating characteristic curves of tuberculous meningitis detected by Xpert MTB/RIF.

Figure 6. Forest plots for the pooled sensitivity of Xpert MTB/RIF.

Figure 7. Forest plots for the pooled specificity of Xpert MTB/RIF.

Figure 8. Forest plots for the pooled positive likelihood ratio of Xpert MTB/RIF.

Figure 9. Forest plots for the pooled negative likelihood ratio of Xpert MTB/RIF.

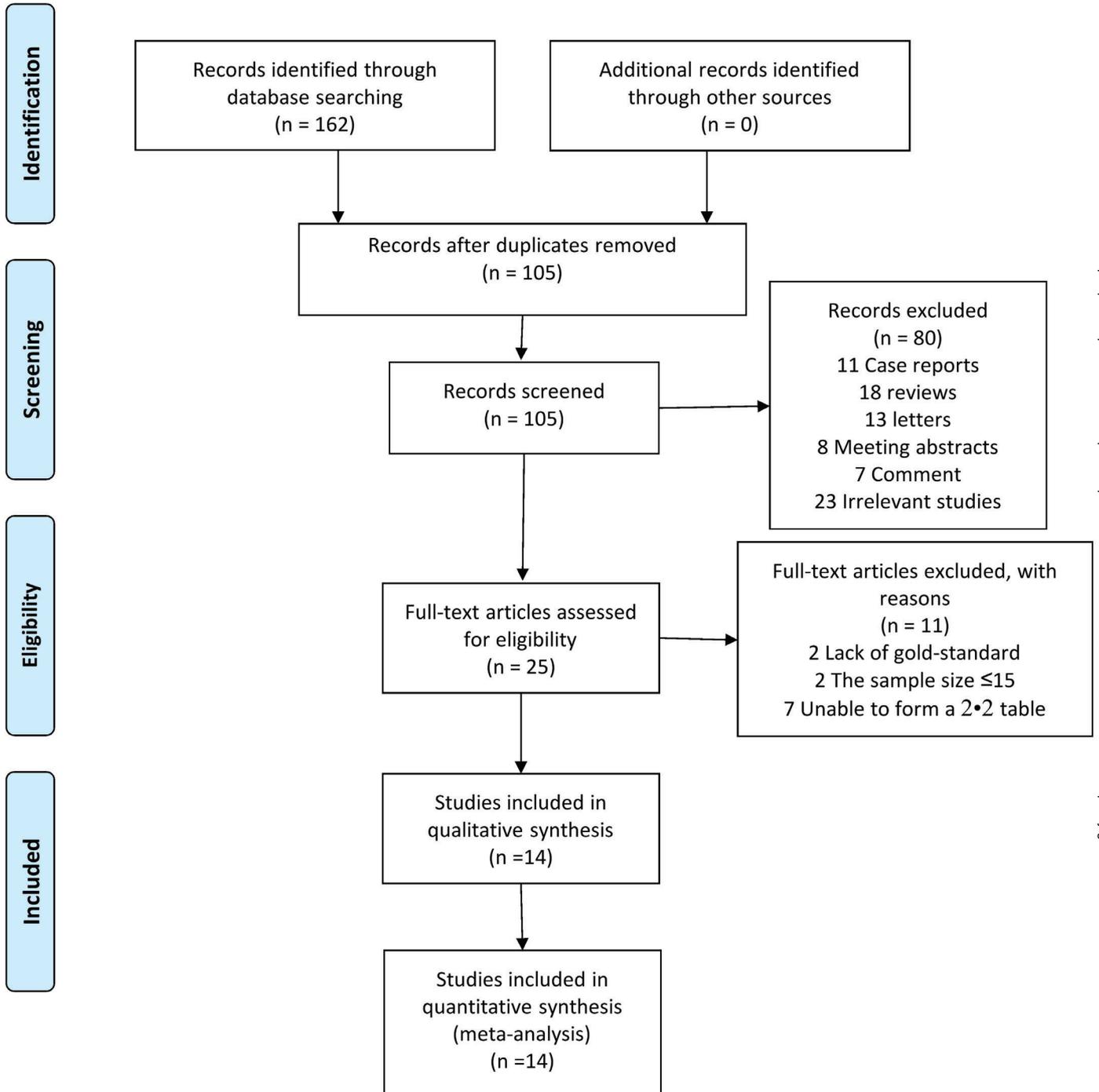
Figure 10. Forest plots for the pooled diagnostic odds ratio of Xpert MTB/RIF.

Table 1. Characteristics of the included studies

Table 2. Subgroup analysis results



## PRISMA 2009 Flow Diagram



From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit [www.prisma-statement.org](http://www.prisma-statement.org).

Risk of Bias

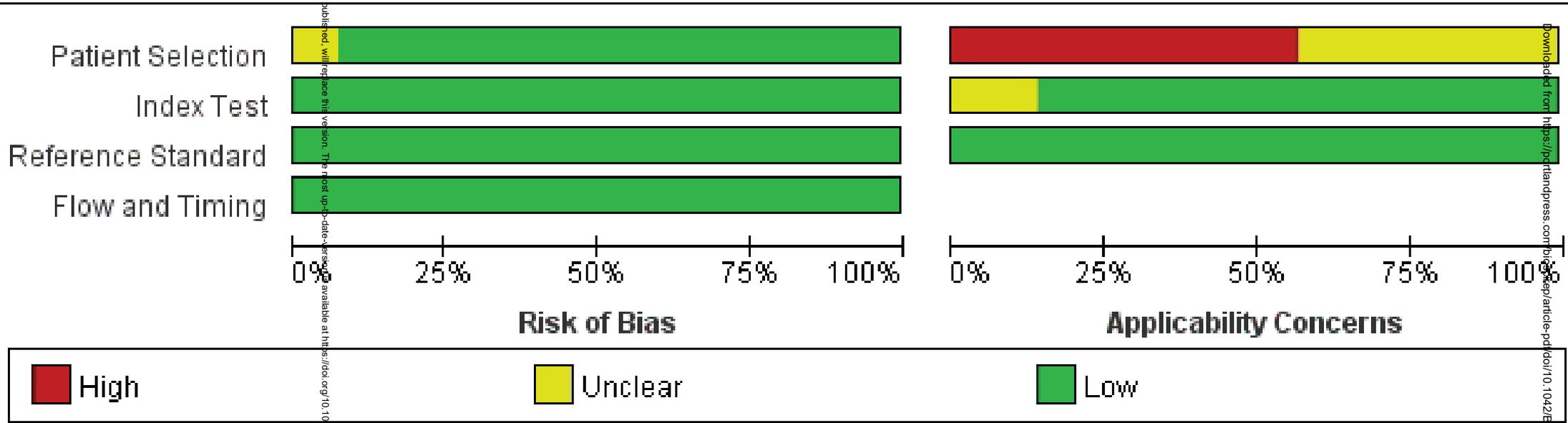
Applicability Concerns

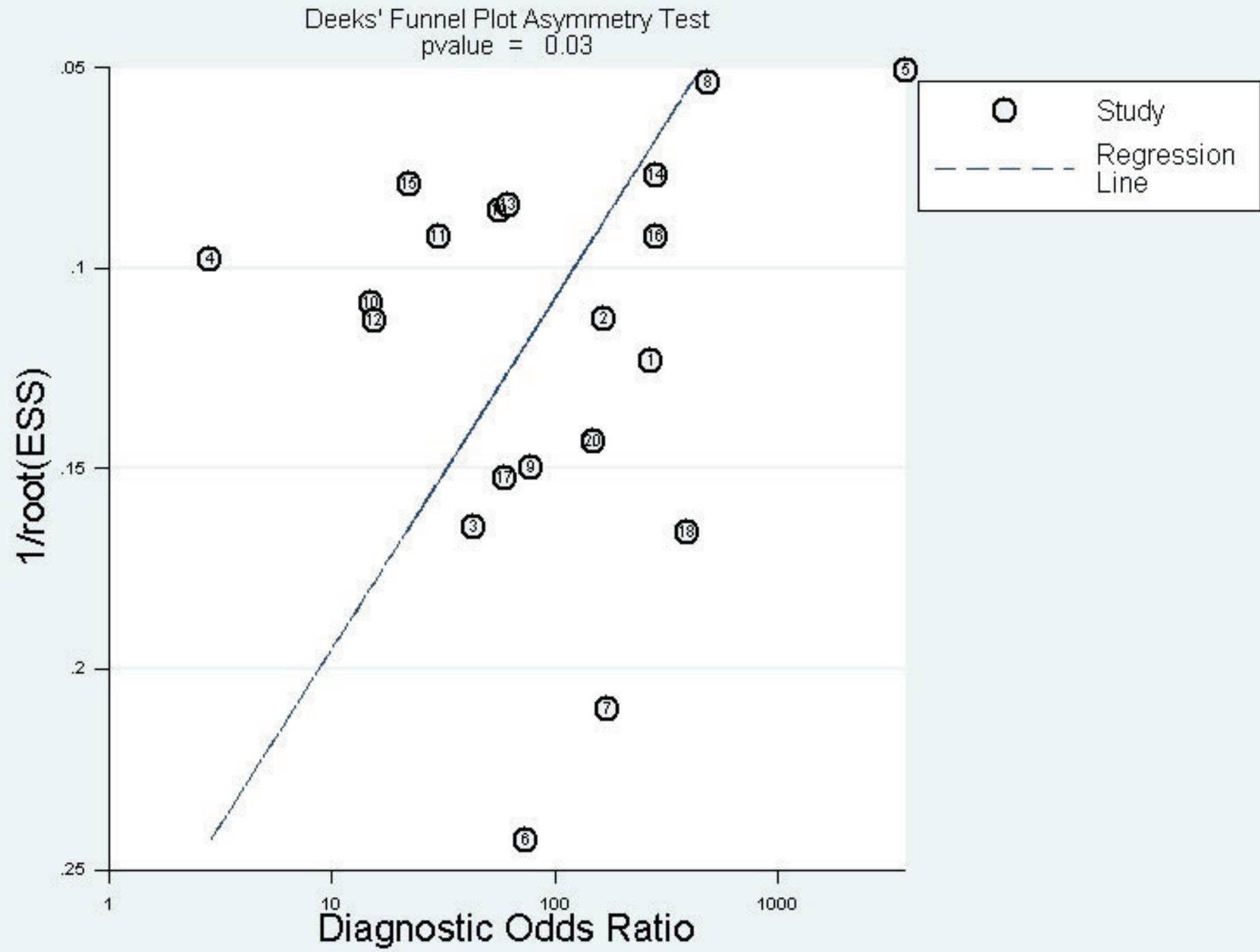
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Bahr 2015	+	+	+	+	-	+	+
Bahr 2017	+	+	+	+	-	+	+
Cresswell 2018	+	+	+	+	-	?	+
Heemskerk 2018	+	+	+	+	-	?	+
Metcalf 2018	+	+	+	+	-	+	+
Nhu 2013	+	+	+	+	?	+	+
Patel 2013	+	+	+	+	?	+	+
Patel 2014	+	+	+	+	-	+	+
Pink 2015	+	+	+	+	?	+	+
Rufai 2017	?	+	+	+	?	+	+
Sharma 2018	+	+	+	+	-	+	+
Solomons 2015	+	+	+	+	-	+	+
Solomons 2016	+	+	+	+	?	+	+
Wang 2016	+	+	+	+	?	+	+

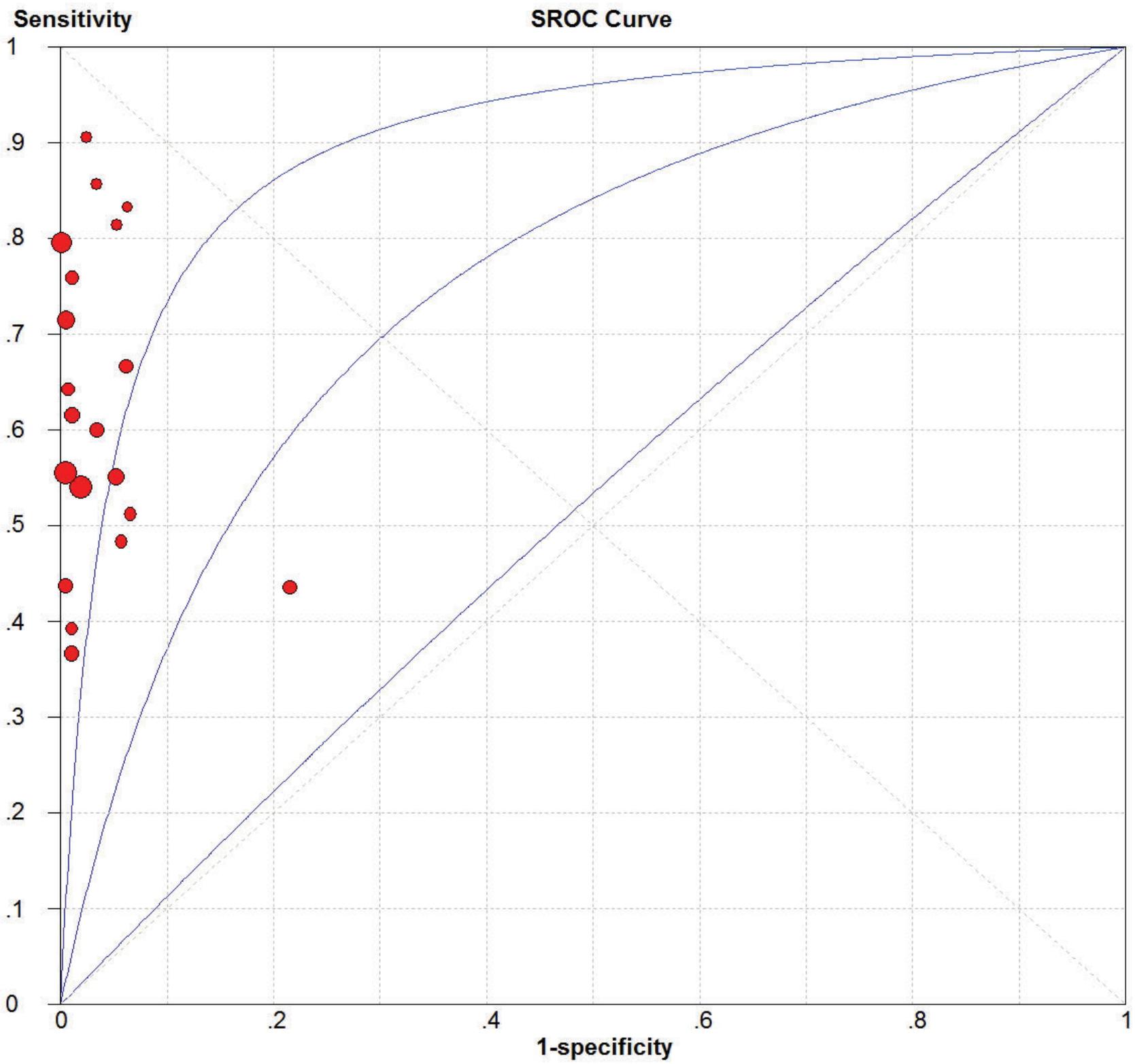
High

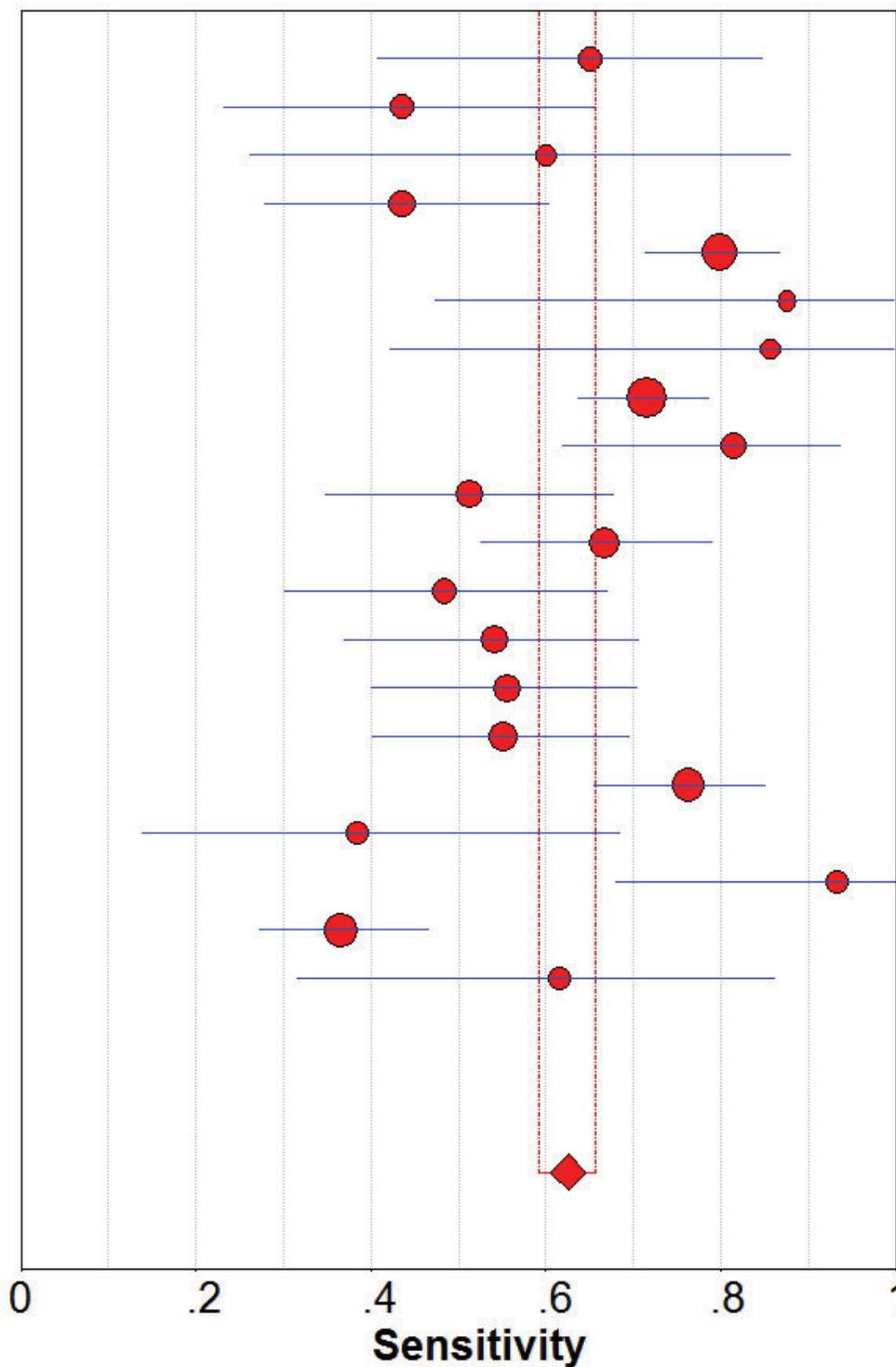
Unclear

Low



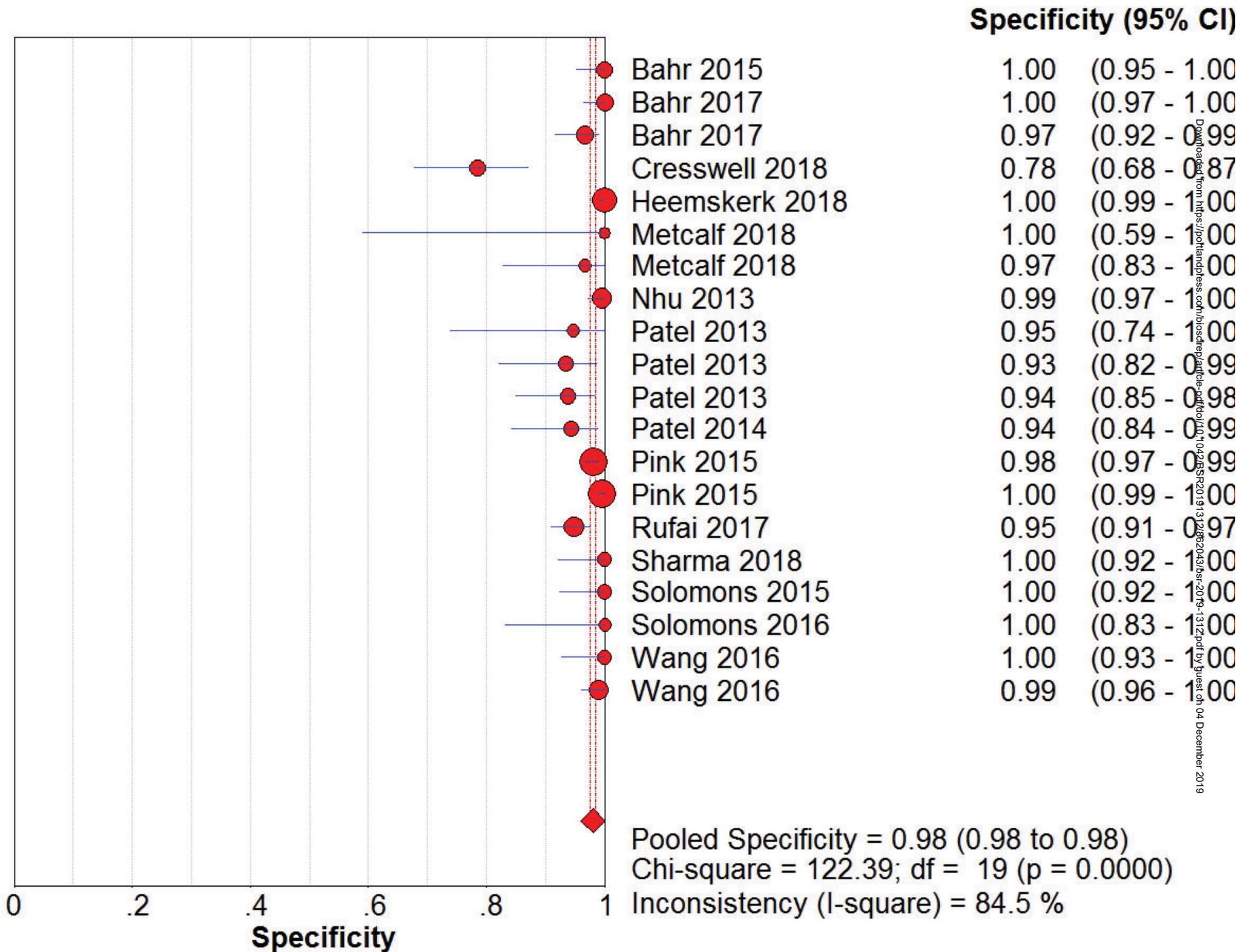


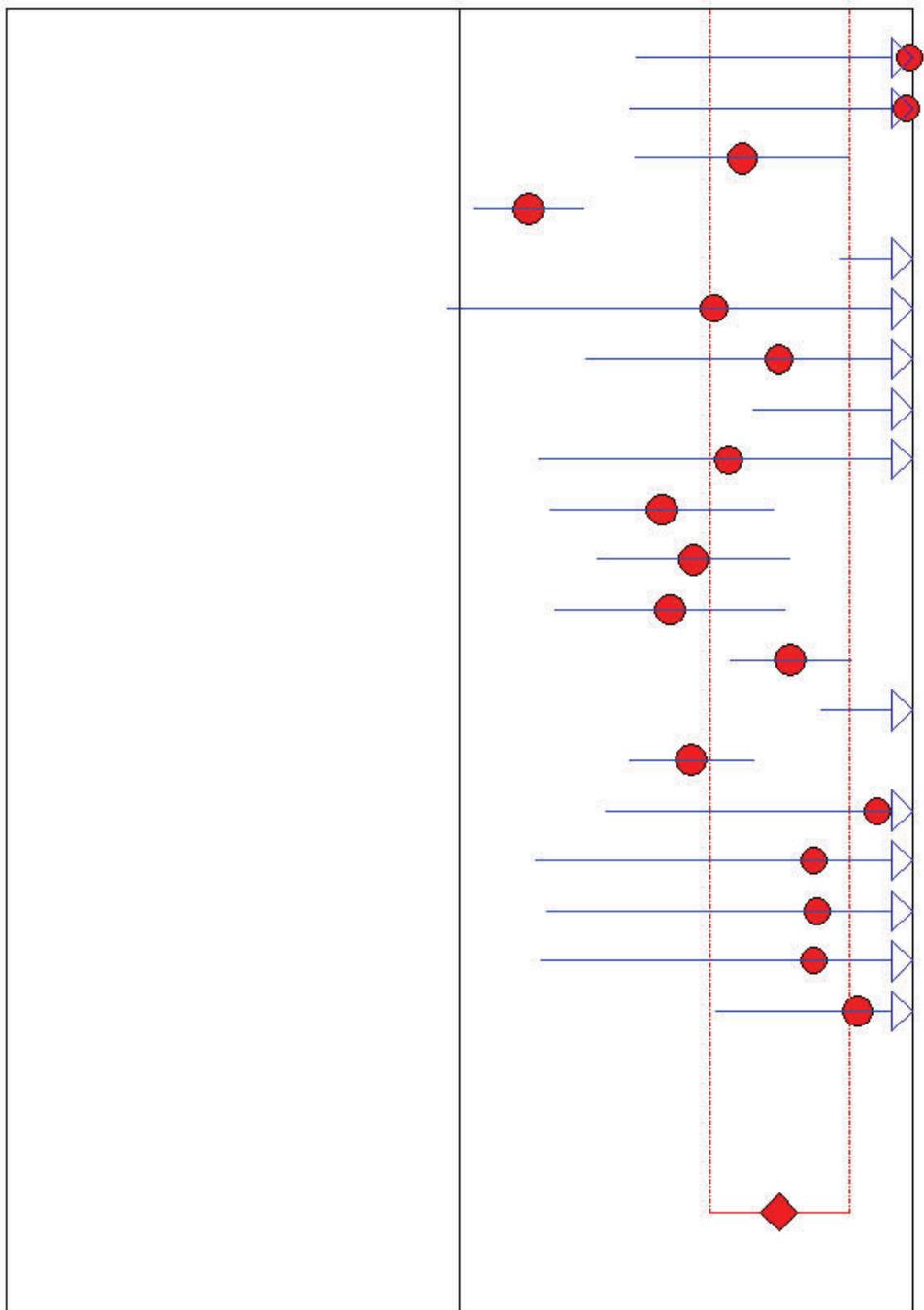




	Sensitivity (95% CI)	
Bahr 2015	0.65	(0.41 - 0.85)
Bahr 2017	0.43	(0.23 - 0.66)
Bahr 2017	0.60	(0.26 - 0.88)
Cresswell 2018	0.44	(0.28 - 0.60)
Heemskerk 2018	0.80	(0.71 - 0.87)
Metcalf 2018	0.88	(0.47 - 1.00)
Metcalf 2018	0.86	(0.42 - 1.00)
Nhu 2013	0.72	(0.64 - 0.79)
Patel 2013	0.81	(0.62 - 0.94)
Patel 2013	0.51	(0.35 - 0.68)
Patel 2013	0.67	(0.53 - 0.79)
Patel 2014	0.48	(0.30 - 0.67)
Pink 2015	0.54	(0.37 - 0.71)
Pink 2015	0.56	(0.40 - 0.70)
Rufai 2017	0.55	(0.40 - 0.69)
Sharma 2018	0.76	(0.65 - 0.85)
Solomons 2015	0.38	(0.14 - 0.68)
Solomons 2016	0.93	(0.68 - 1.00)
Wang 2016	0.37	(0.27 - 0.47)
Wang 2016	0.62	(0.32 - 0.86)

Pooled Sensitivity = 0.63 (0.59 to 0.66)  
 Chi-square = 94.89; df = 19 (p = 0.0000)  
 Inconsistency (I-square) = 80.0 %





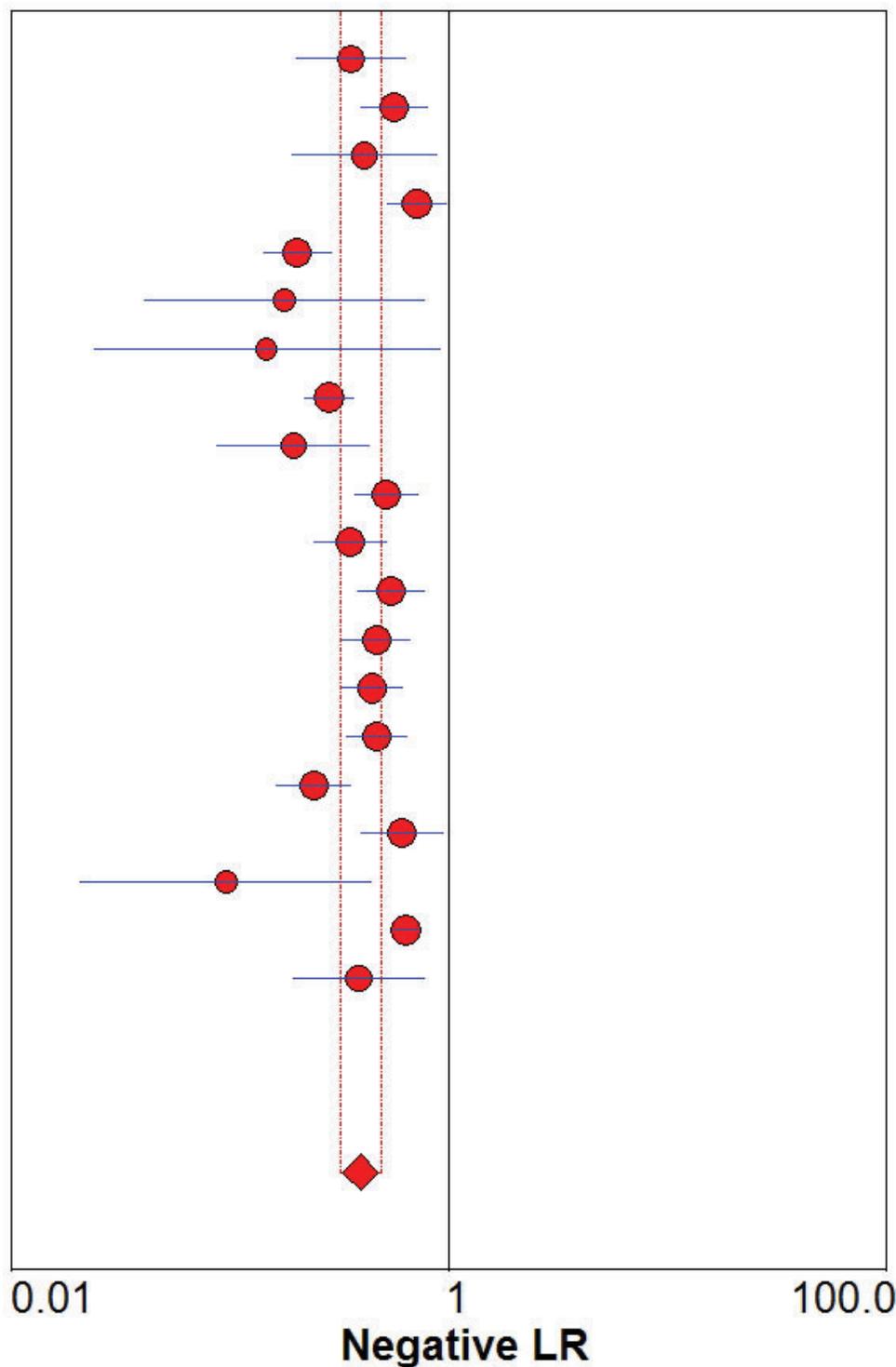
**Positive LR (95% CI)**

Bahr 2015	97.71	(6.06 - 1,576.66)
Bahr 2017	93.63	(5.68 - 1,543.17)
Bahr 2017	17.85	(6.01 - 53.00)
Cresswell 2018	2.03	(1.17 - 3.52)
Heemskerk 2018	770.37	(48.18 - 12,316.96)
Metcalf 2018	13.33	(0.90 - 198.33)
Metcalf 2018	25.71	(3.66 - 180.85)
Nhu 2013	140.90	(19.90 - 997.88)
Patel 2013	15.48	(2.28 - 105.19)
Patel 2013	7.86	(2.52 - 24.49)
Patel 2013	10.83	(4.12 - 28.52)
Patel 2014	8.55	(2.69 - 27.21)
Pink 2015	29.23	(15.80 - 54.07)
Pink 2015	127.78	(40.10 - 407.18)
Rufai 2017	10.62	(5.66 - 19.91)
Sharma 2018	69.85	(4.42 - 1,103.03)
Solomons 2015	36.93	(2.17 - 627.59)
Solomons 2016	38.06	(2.45 - 591.33)
Wang 2016	36.67	(2.30 - 584.77)
Wang 2016	57.85	(13.65 - 245.12)

Random Effects Model  
 Pooled Positive LR = 25.91 (12.71 to 52.82)  
 Cochran-Q = 109.06; df = 19 (p = 0.0000)  
 Inconsistency (I-square) = 82.6 %  
 Tau-squared = 1.7952

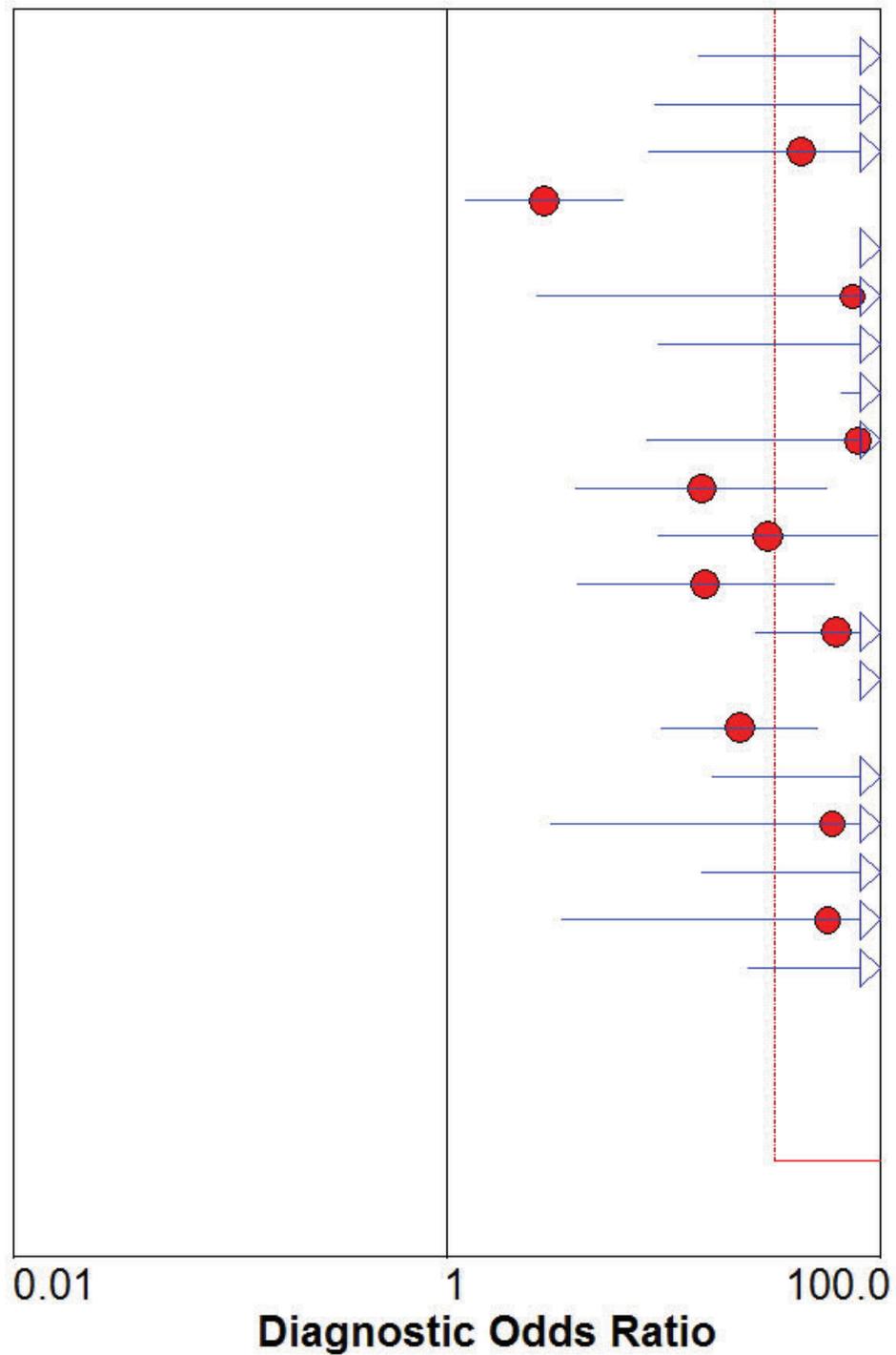
0.01                      1                      100.0

**Positive LR**



	Negative LR	95% C
Bahr 2015	0.36	(0.20 - 0.64)
Bahr 2017	0.57	(0.40 - 0.80)
Bahr 2017	0.41	(0.19 - 0.88)
Cresswell 2018	0.72	(0.53 - 0.97)
Heemskerk 2018	0.20	(0.14 - 0.29)
Metcalf 2018	0.18	(0.04 - 0.77)
Metcalf 2018	0.15	(0.02 - 0.91)
Nhu 2013	0.29	(0.22 - 0.37)
Patel 2013	0.20	(0.09 - 0.43)
Patel 2013	0.52	(0.37 - 0.73)
Patel 2013	0.36	(0.24 - 0.52)
Patel 2014	0.55	(0.39 - 0.77)
Pink 2015	0.47	(0.33 - 0.66)
Pink 2015	0.45	(0.32 - 0.62)
Rufai 2017	0.47	(0.35 - 0.65)
Sharma 2018	0.24	(0.17 - 0.36)
Solomons 2015	0.61	(0.40 - 0.94)
Solomons 2016	0.10	(0.02 - 0.44)
Wang 2016	0.64	(0.55 - 0.74)
Wang 2016	0.39	(0.20 - 0.77)

Random Effects Model  
 Pooled Negative LR = 0.40 (0.32 to 0.50)  
 Cochran-Q = 115.24; df = 19 (p = 0.0000)  
 Inconsistency (I-square) = 83.5 %  
 Tau-squared = 0.1750



### Diagnostic OR (95% CI)

Bahr 2015	271.80	(14.65 - 5,043.14)
Bahr 2017	165.67	(9.18 - 2,989.85)
Bahr 2017	43.13	(8.61 - 215.90)
Cresswell 2018	2.82	(1.23 - 6.46)
Heemskerk 2018	3,769.33	(227.26 - 62,517.23)
Metcalf 2018	75.00	(2.61 - 2,152.37)
Metcalf 2018	174.00	(9.50 - 3,187.27)
Nhu 2013	492.28	(66.86 - 3,624.75)
Patel 2013	79.20	(8.47 - 740.67)
Patel 2013	15.09	(4.00 - 56.94)
Patel 2013	30.50	(9.57 - 97.20)
Patel 2014	15.63	(4.00 - 60.97)
Pink 2015	62.44	(26.74 - 145.80)
Pink 2015	286.25	(79.79 - 1,026.89)
Rufai 2017	22.43	(9.80 - 51.32)
Sharma 2018	287.00	(16.88 - 4,878.79)
Solomons 2015	60.18	(3.04 - 1,191.83)
Solomons 2016	396.33	(15.06 - 10,433.44)
Wang 2016	57.32	(3.44 - 955.74)
Wang 2016	148.80	(24.94 - 887.69)

Random Effects Model  
 Pooled Diagnostic Odds Ratio = 71.49 (32.64 to 156.56)  
 Cochran-Q = 83.86; df = 19 (p = 0.0000)  
 Inconsistency (I-square) = 77.3 %  
 Tau-squared = 2.1034

Author	Year	Country	Reference standard	Cerebrospinal fluid centrifugation	Included patients, n	TP	FP	FN	TN	Prevalence rate (n %)	Prevalence rate $\geq 30\%$
Bahr	2015	Uganda	CRS	YES	95	13	0	7	75	21	NO
Bahr	2017	Uganda	CRS	YES	129	10	0	13	106	18	NO
Bahr	2017	Uganda	Culture	YES	129	6	4	4	115	8	NO
Cresswell	2018	Uganda	Culture	YES	118	17	17	22	62	33	YES
Heemskerk	2018	Vietnam; South Africa; Indonesia	Culture	YES	602	95	0	24	483	20	NO
Metcalfe	2018	Peru	CRS	YES	15	7	0	1	7	53	YES
Metcalfe	2018	Peru	Culture	YES	37	6	1	1	29	19	NO
Nhu	2013	Vietnam	CRS	YES	379	10	1	43	196	40	YES
Patel	2013	South Africa	Culture	YES	46	22	1	5	18	59	YES
Patel	2013	South Africa	Culture	NO	85	20	3	19	43	46	YES
Patel	2013	South Africa	Culture	YES&NO <sup>a</sup>	119	36	4	18	61	45	YES
Patel	2014	South Africa	CRS	NO	84	15	3	16	50	37	YES
Pink	2015	United Kingdom	Culture	YES	740	20	13	17	690	5	NO
Pink	2015	United Kingdom	Culture	YES	735	25	3	20	687	6	NO
Rufai	2017	India	Culture	NO	261	27	11	22	201	19	NO
Sharma	2018	Northern India	Culture	YES	125	61	0	19	45	64	YES
Solomons	2015	South Africa	CRS	YES	59	5	0	8	46	22	NO
Solomons	2016	South Africa	Culture	NO	35	14	0	1	20	43	YES
Wang	2016	China	CRS	YES	153	38	0	66	49	68	YES
Wang	2016	China	Culture	YES	201	8	2	5	186	6	NO

PS: a: indicates that the sample in the data contains samples for centrifugation and non-centrifugation; CRS (Marais, S. et al.): composite reference standard; TP: true positive; FP: false positive; FN: false negative; TN : True negative.

Subgroup analysis	Sensitivity(95%CI)	Specificity(95%CI)	PLR(95%CI)	NLR(95%CI)	DOR(95%CI)	SROC	
						AUC	Q*
<b>GROUP A</b>							
Culture	0.67 (0.63-0.71)	0.98 (0.97-0.98)	22.45 (9.69-52.16)	0.37 (0.28-0.48)	65.73 (25.03-72.62)	0.7204	0.6693
CRS	0.56 (0.51-0.61)	0.99 (0.98-1.00)	35.05 (11.21-109.62)	0.46 (0.32-0.67)	86.47 (25.54-292.75)	0.8592	0.7899
<b>GROUP B</b>							
Yes	0.63 (0.60-0.67)	0.99 (0.98-0.99)	39.11 (13.70-111.64)	0.39 (0.29-0.51)	111.17 (36.80-335.87)	0.7895	0.7267
No	0.57 (0.48-0.65)	0.95 (0.92-0.97)	10.06 (6.17-16.41)	0.49 (0.37-0.64)	21.41 (10.55-43.44)	0.9985	0.9876
<b>GROUP C</b>							
Yes	0.62 (0.57-0.66)	0.95(0.93-0.97)	15.24(4.88-47.55)	0.38 (0.27-0.54)	44.42(12.85-153.56)	0.8352	0.7674
No	0.64 (0.59-0.69)	0.99 (0.98-0.99)	42.54 (18.54-97.63)	0.42 (0.32-0.54)	0.42 (0.32-0.54)	0.6661	0.6262

PS: CI: confidence interval; PLR: positive likelihood ratio; NLR: negative likelihood ratio; DOR: diagnostic odds ratio; SROC: summary receiver operator characteristic curve;AUC: area under the curve; CRS (Marais, S. et al) : composite reference standard.



tuberculosis"[All Fields] OR ("tuberculosis"[All Fields] AND "meningitides"[All Fields])) OR ("tuberculosis, meningeal"[MeSH Terms] OR ("tuberculosis"[All Fields] AND "meningeal"[All Fields]) OR "meningeal tuberculosis"[All Fields] OR ("tuberculosis"[All Fields] AND "hypertrophic"[All Fields] AND "pachymeningitis"[All Fields]) OR "tuberculous hypertrophic pachymeningitis"[All Fields]) OR ("tuberculosis, meningeal"[MeSH Terms] OR ("tuberculosis"[All Fields] AND "meningeal"[All Fields]) OR "meningeal tuberculosis"[All Fields] OR ("hypertrophic"[All Fields] AND "pachymeningitis"[All Fields] AND "tuberculosis"[All Fields])) OR ("tuberculosis, meningeal"[MeSH Terms] OR ("tuberculosis"[All Fields] AND "meningeal"[All Fields]) OR "meningeal tuberculosis"[All Fields] OR ("hypertrophic"[All Fields] AND "pachymeningitis"[All Fields] AND "tuberculous"[All Fields])) OR ("tuberculosis, meningeal"[MeSH Terms] OR ("tuberculosis"[All Fields] AND "meningeal"[All Fields]) OR "meningeal tuberculosis"[All Fields] OR ("pachymeningitis"[All Fields] AND "tuberculous"[All Fields] AND "hypertrophic"[All Fields])) OR ("tuberculosis, meningeal"[MeSH Terms] OR ("tuberculosis"[All Fields] AND "meningeal"[All Fields]) OR "meningeal tuberculosis"[All Fields] OR ("tuberculosis"[All Fields] AND "hypertrophic"[All Fields] AND "pachymeningitis"[All Fields]) OR "tuberculous hypertrophic pachymeningitis"[All Fields])) AND (xpert[All Fields] OR (xpert[All Fields] AND mtb[All Fields]) OR (xpert[All Fields] AND mtb/rif[All Fields]))

**Embase:**

('tuberculous meningitis'/exp OR 'intracranial tuberculoma'/exp) AND ('xpert'/exp OR 'xpert mtb rif'/exp OR 'xpert mtb rif assay'/exp OR 'xpert mtb rif test'/exp)

**WOS:**

(tuberculosis, meningeal OR meningeal tuberculosis OR meningeal tuberculosis OR tubercloses, meningeal tb meningitis or tb meningitides OR tubercular meningitis OR meningitides circular OR meningitis, tubercular OR tubercular meningitides OR meningitis, tuberculous OR meningitides, tuberculous OR tuberculous meningitides OR tuberculous meningitis OR tuberculosis meningitis OR meningitides, tuberculosis OR meningitis, tuberculosis OR tuberculosis meningitides OR tuberculous hypertrophic pachymeningitis OR hypertrophic pachymeningitis, tuberculous OR hypertrophic pachymeningitis, tuberculous OR pachymeningitis, tuberculous hypertrophic OR pachymeningitis, tuberculous hypertrophic OR tuberculous hypertrophic pachymeningitis) AND (xpert OR xpert mtb )