S1_File

Performance of diagnostic and predictive host blood transcriptomic signatures for tuberculosis disease: a systematic review and meta-analysis.

Data Extraction Form

Reviewer initials   □ HM   □ CZZ   □ Other_____

Part A: Study characteristics

Author's last name: _____   **Publication year:** _____

Study title:

_____

Country of study population: _____   **Study #_____**

TB Burden in study population: □Low   □Intermediate □ High

Study design: □Cross-sectional   □Cohort   □ Case control   □ RCT   □ Other _____

Study type: □ Diagnostic   □ Predictive   *if predictive; follow-up time_____months*

Study purpose: □ Discovery   □ Validation

Sampling: □ Consecutive   □Convenient   □Random   □Other_____

Part C: Characteristics of tests

Index sample type   **:** □ Whole blood   □ PBMC

Index test (signature) type : □mRNA   □Other_____

Signature name   **:** _____   **# of genes:** _____

Signature discovery method: □ RNA Seq   □Micro array   □PCR   □Other_____

Signature model   **:** □ Pairwise   □Random forest □SVM   □Other_____

TB disease gold standard : □Culture   □Xpert MTB/RIF □ Smear   □ Other_____

Part B: Population characteristics

Population: □Adults   □Adolescents   □ Children   □Mixed   □ Undefined

Age range: _____

Participant's cohort (test/validation) and disease status**

Cohort Type	Healthy Controls[a]	LTBI[b]	Other Diseases[c]	TB Disease[d]	Total Enrolled

** For diagnostic studies, the number of participants in each category represents the numbers at enrolment while for predictive studies; a, b and c represent the number of participants in each category at enrolment that gave rise to the cases (d) at the end of follow-up period.

S2_File

S1_File

Description of negative population

[a]Healthy Controls: □ Healthy endemics □ Healthy non-endemics □ TB Contacts □ Other, describe:_____

_

[b]Other diseases: □ Non-respiratory diseases □ Respiratory diseases

[c]LTBI : □TST>5mm □TST>10mm □IGRA □ N/A

If IGRA, IGRA type : □QFT □Elispot □Other_____

[d]*TB disease: For diagnostic studies, this number represents the number of participants included in the study at enrolment while for predictive studies; this number represents the number of participants diagnosed with TB at the end of follow-up period.*

Part D: Outcomes

Cohort Type	Sens	Spec	AUC	TP	TN	FP	FN	LR+	LR-	RR

Key: *Sens= Sensitivity; Spec= Specificity; AUC= Area under the curve; TP = True positives; FP = False positives; TN = True negatives; FN = False negatives; LR+ = Positive likelihood ratio; LR- = Negative likelihood ratio; RR = Rate Ratio for predictive studies only.*

Inclusion/Exclusion Decision

□Included □Excluded □ Pending

If excluded, reason:

□ Excluded by title

□ Children only

□ Different index test

□ Reference test neither Culture, Xpert/MTB RIF nor Smear Microscopy

□ HIV infected only

□ Mixed HIV infected and uninfected and could not desegregate the participants

□ Mixed children and adults and could not desegregate the participants

□ Other study design

□ Diagnostic performance data not reported and unable to get from authors

S2_File