Single-gene transcripts for subclinical tuberculosis: an individual participant data meta-analysis

James Greenan-Barrett, Simon C Mendelsohn, Thomas J Scriba, Mahdad Noursadeghi, Rishi K Gupta**

Summary

Background Translation of blood RNA signatures might be accelerated by identifying biomarkers composed of the minimum number of gene transcripts. We aimed to test the hypothesis that single-gene transcripts provide similar accuracy for detection of subclinical tuberculosis to multi-gene signatures and benchmark their accuracy and clinical utility against interferon-γ release assays (IGRAs).

Methods For this individual participant data meta-analysis, we searched PubMed from database inception to June 10, 2024, using terms for "tuberculosis", "subclinical", and "RNA" to identify studies in which participants underwent whole-blood RNA sampling with at least 12 months of follow-up for development of clinical tuberculosis. We performed a one-stage individual participant data meta-analysis to compare the accuracy of multi-gene signatures against single-gene transcripts to discriminate individuals with subclinical tuberculosis—defined as asymptomatic prevalent or incident tuberculosis (diagnosed ≥21 days from enrolment, irrespective of symptoms) over a 12-month interval—from individuals who remained disease free. We performed decision curve analysis to evaluate the net benefit of using single-gene transcripts and IGRAs, alone or in combination, to stratify preventive treatment compared with strategies of treating all or no individuals.

Findings 276 articles were identified in the search; of these, seven met the eligibility criteria and all had IPD available. We evaluated 80 single-genes and eight multi-gene signatures in a pooled analysis of four RNA sequencing and three quantitative PCR datasets, comprising 6544 total samples and including 283 samples from 214 individuals with subclinical tuberculosis. Distributions of transcript and signature Z scores after standardisation were similar and there was little heterogeneity between datasets. Five single-gene transcripts (*BATF2*, *FCGR1A/B*, *ANKRD22*, *GBP2*, and *SERPING1*) had equivalent areas under the receiver operating characteristic curves (0·75 [95% CI 0·71–0·79] to 0·77 [0·73–0·81]) to the best-performing multi-gene signature over 12 months, but none met the WHO minimum target product profile (TPP) for a tuberculosis progression test. IGRAs approximated the TPP in low-burden settings but showed much lower specificity in high-burden settings (74% [95% CI 72–76] *vs* 32% [30–35]). By contrast, sensitivity (67% [47–82] in high-burden settings *vs* 78% [67–86] in low-burden settings) and specificity (72% [70–74] *vs* 67% [64–69]) of the best-performing single-gene transcript was similar across settings. Decision curve analysis showed that in high-burden settings, stratifying preventive treatment using single-gene transcripts had greater net benefit than using IGRAs, which offered little net benefit over treating all individuals. In low-burden settings, IGRAs offered greater net benefit than single-gene transcripts to stratify treatment, but combining both tests provided the highest net benefit for tuberculosis programmes aiming to treat fewer than 50 people to prevent a single case.

Interpretation Single-gene transcripts are equivalent to multi-gene signatures for detection of subclinical tuberculosis, with consistent performance across settings. Single-gene transcripts show potential clinical utility to stratify preventive treatment, particularly when used in combination with IGRAs in low-burden settings.

Funding National Institute for Health Research, Wellcome Trust.

Lancet Microbe 2025

Published Online
https://doi.org/10.1016/j.lanmic.2025.101186

*Co-senior authors

UCL Respiratory, Division of Medicine, University College London, London, UK (J Greenan-Barrett MRCP, R K Gupta PhD); **South African Tuberculosis Vaccine Initiative, Division of Immunology, Department of Pathology and Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Cape Town, South Africa** (S C Mendelsohn PhD, Prof T J Scriba PhD); **UCL Division of Infection and Immunity, University College London, London, UK** (Prof M Noursadeghi PhD)

Correspondence to:
Dr Rishi K Gupta, UCL Respiratory, Division of Medicine, University College London, London WC1E 6JF, UK
r.gupta@ucl.ac.uk

Introduction

Despite global efforts, tuberculosis is a leading cause of morbidity and mortality worldwide, with 10·6 million cases and 1·3 million deaths in 2022, and a disproportionate burden among disadvantaged communities.[1,2] In response to increased recognition of the spectrum of tuberculosis, the International Consensus for Early Tuberculosis (ICE-TB) group developed a framework to classify disease states in 2024. The ICE-TB framework divides tuberculosis disease into subclinical or clinical based on signs and symptoms, with further subdivisions of each into infectious and non-infectious based on the detection of aerosolised or expectorated *Mycobacterium tuberculosis*.[3] Targeting the subclinical disease state, potentially with truncated treatment regimens,[4] could prevent disease progression and reduce the risk of onward transmission.

Tests of immunoreactivity to *M tuberculosis* used to stratify tuberculosis preventive therapy, such as the

Research in context

Evidence before this study

With increasing recognition of the spectrum of tuberculosis and the importance of subclinical tuberculosis, in 2024 the International Consensus for Early Tuberculosis (ICE-TB) group developed a framework to classify disease states and called for diagnostic development to detect the subclinical states. We performed a systematic search of PubMed from database inception to June 10, 2024, using the terms "tuberculosis", "subclinical", and "RNA", without language restrictions. Multiple studies have discovered or validated multi-gene blood RNA signatures for tuberculosis, including to differentiate individuals who progress to clinical tuberculosis from non-progressors. In our previous head-to-head evaluation of these signatures, we showed that eight signatures had equivalent performance to predict progression to tuberculosis. Simplifying biomarkers to single-gene transcripts could facilitate clinical translation. However, no previous studies have systematically compared the performance of single-gene transcripts to multi-gene signatures. Moreover, the clinical utility of RNA biomarkers to detect subclinical tuberculosis and guide preventive treatment decisions, compared with alternative strategies using interferon-γ release assays (IGRAs), is untested.

Added value of this study

To our knowledge, this study is the largest pooled analysis of RNA biomarkers to predict progression to clinical tuberculosis, the first RNA analysis to align with the ICE-TB definitions, and the first head-to-head comparison of single-gene transcripts with multi-gene signatures. We tested 80 single-genes and eight multi-gene signatures to detect subclinical tuberculosis in a pooled dataset from four RNA sequencing and three quantitative PCR datasets, comprising over 6500 RNA samples. We showed that five co-correlated single-gene transcripts were equivalent to the best-performing multi-gene signature to detect subclinical tuberculosis over a 12-month interval, but none met the WHO minimum target product profile for a tuberculosis progression test. We showed that single-gene transcripts performed consistently across settings, whereas IGRA performance was heterogeneous, with poor specificity in high-burden settings. By using decision curve analysis, we showed that single-gene transcripts offered highest net benefit in high-burden settings, whereas IGRA offered highest net benefit in low-burden settings, either alone or in combination with single-gene transcripts in a two-step approach.

Implications of all the available evidence

Single-gene transcripts perform as well as multi-gene signatures to diagnose subclinical tuberculosis. These findings might simplify RNA biomarker testing, encourage commercial competition, and facilitate translation of this technology into clinical practice. Blood RNA biomarkers showed clinical utility to direct preventive treatment decisions as a stand-alone test in high-burden settings and in combination with IGRA in low-burden settings. Further interventional studies are required to evaluate the clinical and cost-effectiveness of serial RNA biomarker testing to stratify delivery of preventive treatment among individuals at high risk in high-burden countries or using a two-step testing approach in combination with IGRA in low-burden settings.

tuberculin skin test and interferon-γ release assay (IGRA), have low positive predictive values (PPVs) for progression to clinical disease,[5–7] resulting in unnecessary treatment for most individuals. Preventive therapy can be financially burdensome on health-care systems and lead to adverse effects in 3·7% of individuals.[8]

Measurement of blood RNA can be used to detect changes in host gene expression in response to tuberculosis disease. Multiple RNA signatures have been discovered that have promising diagnostic accuracy for clinical tuberculosis[9–20] or to predict progression to clinical disease.[21–24] In a previous analysis comparing the performance of 17 RNA signatures to predict progression to clinical disease, we showed that eight signatures performed equivalently, were co-correlated, and shared common upstream pathways.[25]

The development of near-patient, cartridge-based prototype platforms, such as the Xpert MTB Host Response (MTB-HR; Cepheid, Sunnydale, CA, USA) assay, have advanced progress towards clinical translation of blood RNA signatures.[26,27] However, this progress might be further accelerated by simplification of multi-gene signatures to single-gene biomarkers. In view of the co-correlation and common regulators of the genes that comprise the most accurate signatures to date, multiple genes might not offer orthogonal value. In this study, we used a pooled dataset of studies with blood RNA sampling and longitudinal follow-up for tuberculosis to hypothesise that single-gene transcripts will perform as well as multi-gene signatures for subclinical tuberculosis detection. We also sought to benchmark the diagnostic performance and clinical utility of RNA biomarkers to IGRAs, a widely available prognostic test, and stratify these analyses by tuberculosis burden.

See Online for appendix

Methods

Search strategy and selection criteria

For this one-stage individual participant data (IPD) meta-analysis, we performed a systematic search of PubMed from database inception to June 10, 2024, using terms for "tuberculosis", "subclinical", and "RNA" to identify studies in which participants underwent whole-blood RNA sampling with at least 12 months of follow-up for development of clinical tuberculosis. Studies using genome-wide (RNA sequencing or microarray) or targeted transcriptional profiling (quantitative PCR [qPCR] or NanoString quantification) were included. Full details on the search, including search terms, are provided in the appendix (p 3). JG-B performed the search and any uncertainties about

eligibility were discussed and resolved between JG-B, MN, and RKG. We included four studies using RNA sequencing from our previous IPD meta-analysis[25] and three studies using qPCR identified from the systematic search.[28–30] All included data were publicly available. Since this was a secondary analysis of publicly available data, ethical approval was not required.

Data analysis

Data from four previously included RNA sequencing studies were mapped, batch corrected, and integrated into a single pooled dataset using transcripts per million measurements, as previously described.[25] Our data preparation pipeline for qPCR studies is described in the appendix (pp 3–21). The approach to combining studies is described in the appendix (p 4). We included 80 single-genes that were present in the RNA sequencing dataset and at least one qPCR study (Correlate of Risk Targeted Intervention Study [CORTIS-01],[28] CORTIS-HR,[29] or Regional Prospective Observational Research for Tuberculosis Brazil [REPORT-Brazil][30]). We calculated scores for eight existing RNA signatures that were included in our previous analysis (appendix pp 3–4).[31] We standardised signatures and single-gene transcripts within each RNA sequencing and qPCR dataset by converting to Z scores (appendix pp 14–15), before combining into a pooled dataset.

We used the original study definitions for tuberculosis in each cohort. As baseline screening for prevalent disease was performed variably between studies, we arbitrarily defined prevalent tuberculosis as a tuberculosis diagnosis up to 21 days after RNA sampling.[25,32] Cases diagnosed after 21 days were considered to be incident tuberculosis.

As our aim was to evaluate the accuracy of RNA biomarkers to discriminate between individuals with subclinical tuberculosis and those without disease, we used the ICE-TB consensus classification to define disease states at the point of RNA sampling.[3] Individuals with symptomatic prevalent tuberculosis disease were classified as having clinical tuberculosis at the point of RNA sampling. Individuals with asymptomatic prevalent disease or those who progressed to incident disease were classified as having subclinical tuberculosis at the point of RNA sampling. Those without prevalent or incident disease were defined as being disease free. We approximated asymptomatic prevalent tuberculosis to the subclinical, infectious tuberculosis state and incident tuberculosis to the subclinical, non-infectious tuberculosis state. The subclinical, non-infectious ICE-TB state is based on the presence of macroscopic pathology, for which there is no gold standard. We therefore used progression to incident tuberculosis as a reference standard to approximate this state, based on the assumption that macroscopic pathology would have been detectable at the time of blood RNA sampling if investigated with sufficient resolution.

All analyses were done with R (version 4.4.0). We performed a one-stage IPD meta-analysis to calculate the accuracy of candidate signatures and transcripts to discriminate individuals with subclinical tuberculosis from individuals who remained disease free, stratified by interval from sampling to disease. Initial analyses showed similar accuracy of RNA biomarkers in each study, so our primary analysis assumed common accuracy across studies, as previously described.[25] The primary analysis was over a 12-month interval from sampling, with secondary analyses over months 0–3, 0–6, 0–15, 6–12, and 12–15. For each analysis, disease free samples with follow-up less than the stated interval were excluded to reduce the risk of outcome misclassification. We included participants irrespective of IGRA status but excluded individuals who received tuberculosis preventive therapy from the primary analysis as this affects progression risk and might be differential among those with high RNA biomarker scores. If datasets included serial RNA samples from the same individuals, we considered serial samples as independent because intra-individual variance was similar to inter-individual variance (appendix p 22). If candidate signatures were originally derived from included datasets, we excluded these datasets when evaluating the accuracy of that signature.

Accuracy of candidate signatures and transcripts was quantified by the area under the receiver operating curve (AUROC) with 95% CIs. Sensitivity and specificity were calculated at the maximum Youden index, giving equal weighting to sensitivity and specificity, and benchmarked against the WHO minimum target product profile (TPP) parameters for predicting progression to tuberculosis over 2 years (≥75% sensitivity and ≥75% specificity).[33] AUROCs of single-gene transcripts were compared with the best-performing multi-gene signature using the pairwise Delong test, with multiple testing correction using the Benjamini–Hochberg approach. Signatures and transcripts with adjusted p values over 0·05 were considered equivalent. Correlation of equivalent transcripts was assessed using Spearman's rank correlation. We also evaluated expression and AUROCs of the best-performing single-gene transcript across different disease states: clinical tuberculosis; subclinical tuberculosis, infectious; and subclinical tuberculosis, non-infectious.

We then compared the diagnostic performance (to discriminate individuals with subclinical tuberculosis from individuals who remained disease free) of the single-gene transcript with the highest AUROC point estimate against IGRA in a head-to-head analysis among participants for whom results of both tests were available over a 12-month interval from sampling. These analyses were stratified by setting, defined as low tuberculosis burden if the incidence was less than 50 per 100 000 people and high tuberculosis burden if it was more than 50 per 100 000 people. Among individuals with serial IGRA samples, low intra-individual variance indicated high correlation between serial samples (appendix p 22); therefore, we included only one sample per individual, sampled at random. We used thresholds of the maximum Youden index for the single-gene transcript and the standard cutoff

of 0·35 IU/mL for the IGRA assay (QuantiFERON-TB Gold-in-tube or QuantiFERON-TB Gold Plus, Qiagen). We also explored a two-step approach using a combination of the single-gene transcript and IGRA, in which only those positive for both tests are offered treatment and compared the sensitivity, specificity, and PPVs of these approaches. As datasets included case–control studies, we calculated PPV using a fixed cumulative disease risk (reflecting prior probability) of 1% and 2%, with 2% risk based on 1-year incidence rates in high-risk close contacts.[34,35]

To evaluate the clinical utility of the single-gene transcript, IGRA, and combined testing approaches, we performed decision curve analysis. Decision curve analysis quantifies the trade-off between correctly identifying true tuberculosis cases and incorrectly identifying false positives as net benefit.[36] Net benefit is calculated across a range of weightings for the false positives, defined as the threshold probability (appendix p 4). Threshold probability is the risk of disease at which a clinician or patient would opt for an intervention such as treatment and relates to the number-willing-to-treat (NWT), defined as the number of individuals that a clinician would be willing to treat with preventive therapy to prevent a single case of tuberculosis disease. We calculated net benefit using the best-performing single-gene transcript or IGRA to guide preventive treatment compared with the default strategies of treating all individuals or no individuals across a range of threshold probabilities. Since the contributing datasets included case–control analyses, we fixed the cumulative tuberculosis risk as 1% and 2% in our decision curve analyses.

We also estimated the number-needed-to-treat (NNT), defined as the number of people needing to be treated with preventive therapy to prevent a single case of tuberculosis disease, using the single-gene transcript, IGRA, and combined testing approaches to stratify treatment, and compared with a default strategy of treating all individuals. Here, we assumed a constant treatment effect of 80% and fixed the cumulative tuberculosis risk as 1% and 2%.

We performed four sensitivity analyses. First, we performed a two-stage IPD meta-analysis in which we calculated accuracy of signatures and transcripts for each contributing cohort, explored between study heterogeneity by visualising forest plots, and performed a random-effects meta-analysis to calculate pooled AUROCs (using the metafor package in R).[37] Second, we included participants commencing tuberculosis preventive therapy in the analysis. Third, we included only one RNA sample per individual, sampled at random. Finally, we included datasets from which signatures were originally derived in the accuracy calculation for that signature. As all data were publicly available, this study did not need to be registered.

Role of the funding source

The funders of the study had no role in study design, data collection, data analysis, data interpretation, writing of the report, or decision to submit for publication.

Results

276 articles were identified in the search; of these, seven met the eligibility criteria and all had IPD available. Four RNA sequencing datasets and three qPCR datasets were included. Study characteristics are provided in table 1. The RNA sequencing datasets included the Adolescent Cohort Study of individuals from South Africa with a positive IGRA or tuberculin skin test;[21] the Grand Challenges 6–74 study of tuberculosis household contacts from South Africa, The Gambia, and Ethiopia;[22] and two UK close contact studies from London[24] and Leicester.[38] The qPCR datasets included the CORTIS-01 study of healthy volunteers from tuberculosis endemic communities in South Africa,[28] the CORTIS-HR study consisting of people living with HIV from tuberculosis endemic communities in South Africa,[29] and the REPORT-Brazil study of tuberculosis close contacts from Brazil.[30]

In total, 6530 samples from 5185 individuals were included in the primary analysis, with a total of 283 samples from individuals with subclinical tuberculosis: 39 from asymptomatic prevalent cases and 244 from incident cases (appendix p 23). Baseline participant characteristics and risk of tuberculosis disease for each study are shown in the appendix (p 24).

Three signatures (Penn-Nicholson6, Darboe11, and Suliman4) were derived from included studies (table 1), so the original study was excluded from the evaluation of the derived signature in the primary analysis. Distributions of Z scores after standardisation were similar between datasets and there was little heterogeneity in AUROCs of transcripts between datasets (appendix pp 15, 32). Therefore, we proceeded to a one-stage IPD meta-analysis as the primary analysis.

The multi-gene signature with the highest AUROC for discrimination of individuals with subclinical tuberculosis from individuals who remained disease free over 12 months from sampling was Roe3 (AUROC 0·77 [95% CI 0·73–0·81]; appendix p 30), which was then used for pairwise comparison with the single-gene transcripts. Five single-gene transcripts had equivalent AUROCs to Roe3; *BATF2* (0·77 [0·73–0·81]), *FCGR1A/B* (0·77 [0·73–0·81]), *ANKRD22* (0·77 [0·72–0·81]), *GBP2* (0·75 [0·71–0·79]), and *SERPING1* (0·75 [0·71–0·79]; table 2). These transcripts showed moderate–strong correlation (0·54–0·82) using Spearman's rank correlation (appendix p 25). Using maximum Youden index thresholds, none of the transcripts met the WHO minimum TPP parameters for progression tests over 12 months. Discriminative performance of single-gene transcripts and Roe3 diminished gradually over increasing time intervals from sampling to disease, with poor discrimination over 12–15 months (appendix pp 26–30). *BATF2* expression and discriminative performance across different disease and disease free states is shown in figure 1. Notably, discriminative performance of *BATF2* was highest for clinical tuberculosis (AUROC 0·93 [95% CI 0·87–0·99]), compared with other states.

Comparison of the diagnostic performance of *BATF2*, IGRA, and a combined two-step approach of *BATF2* and

	ACS[21]	CORTIS-01[28]	CORTIS-HR[29]	GC6–74[22]	Leicester Contacts[38]	London Contacts[24]	REPORT-Brazil[30]
Method of RNA profiling	RNA sequencing	qPCR	qPCR	RNA sequencing	RNA sequencing	RNA sequencing	qPCR
Study design	Nested case–control	Randomised controlled trial*	Cohort	Nested case–control	Cohort	Cohort	Cohort
Newcastle–Ottawa Scale score	9/9	7/7	7/7	9/9	7/7	7/7	7/7
Population	Children aged 12–18 years with a positive interferon-γ release assay or tuberculin skin test	Adults aged 18 years or older in an endemic community	Adults aged 18 years or older with HIV in an endemic community	People aged 10–60 years who have household contact with a tuberculosis index case	People aged 16 years or older who have close contact with a tuberculosis index case	People aged 18 years or older who have close contact with a tuberculosis index case	People aged 18 years or older who have close contact with a tuberculosis index case
Location (burden of tuberculosis in setting)	South Africa (high burden)	South Africa (high burden)	South Africa (high burden)	South Africa, The Gambia, Ethiopia (high burden)	UK (low burden)	UK (low burden)	Brazil (low burden)
Baseline screening	Clinical evaluation	Clinical evaluation and two spontaneous sputum Xpert MTB/RIF tests	Clinical evaluation and two spontaneous sputum Xpert MTB/RIF tests	Clinical evaluation	Clinical evaluation and chest x-ray	Clinical evaluation and chest x-ray	Clinical evaluation
Study case definition	Two positive smear or one positive culture	Two positive Xpert MTB/RIF or Xpert Ultra or culture	Two positive Xpert MTB/RIF or Xpert Ultra or culture	Positive culture or clinically diagnosed	Positive culture or Xpert MTB/RIF	Positive culture or clinically diagnosed	Positive culture or clinically diagnosed
Duration of follow-up, months	24	15	15	24	24	22·8	24
RNA sample timing	Baseline and months 6, 12, 18, and 24	Baseline	Baseline	Baseline and months 6 and 18	Baseline	Baseline	Baseline and month 6
Signatures derived from dataset	Penn-Nicholson6 and Darboe11	NA	NA	Suliman4	NA	NA	NA
Total participants	144	2496	404	334	104	324	1379
Total RNA samples	318	2496	404	412	104	324	2472

ACS=Adolescent Cohort Study. CORTIS=Correlate of Risk Targeted Intervention Study. GC6-74=Grand Challenges 6-74. NA=not applicable. qPCR=quantitative PCR. REPORT-Brazil=Regional Prospective Observational Research for Tuberculisis Brazil. *CORTIS-01 was a randomised controlled trial in which a cohort of healthy participants first underwent measurement of the previously identified RISK11 signature (Darboe11). RISK11-positive individuals were randomly assigned to either receive tuberculosis preventive therapy or not, and RISK11-negative individuals did not receive tuberculosis preventive therapy. Only a randomly sampled subset of RISK11-negative participants was included to enrich the study cohort for RISK11-positive participants. In this meta-analysis, only individuals who did not receive tuberculosis preventive therapy were included.

Table 1: **Characteristics of studies included in meta-analysis**

IGRA to discriminate individuals with subclinical tuberculosis from individuals who remained disease free, stratified by setting and benchmarked against the WHO minimum TPP criteria (figure 2A), showed that although the sensitivity of IGRA was similar across settings (87–88%), specificity was markedly lower in high-burden settings (32% [95% CI 30–35]) compared with low-burden settings (74% [72–76]), resulting in a PPV of 1·3% (95% CI 1·1–1·4) for high-burden settings and 3·3% (2·4–3·9) for low-burden settings (appendix p 33). By contrast, *BATF2* performance—sensitivity (67% [95% CI 47–82] in high-burden settings *vs* 78% [67–86] in low-burden settings), specificity (72% [70–74] *vs* 67% [64–49]), and PPVs (2·3% [1·8–2·7] *vs* 2·4% [1·6–3·1])—was consistent across settings. In high-burden settings, the combined approach resulted in a slight increase in specificity compared with *BATF2* alone (77% [75–79] *vs* 67% [64–69]) but similar PPVs (2·8% [2·1–3·5] *vs* 2·3% [1·8–2·7]). By contrast, in low-burden settings using the combined approach compared with using *BATF2* alone, a large increase in specificity (92% [90–93] *vs* 72% [70–74]) and PPV (6·8% [3·8–9·9] *vs* 2·4% [1·6–3·1]) was reached, albeit with some loss of sensitivity (58% [39–76] *vs* 67% [47–82]). None of the testing approaches met the WHO minimum TPP parameters over 12 months, although IGRA accuracy approached the benchmarks in the low-burden settings. The testing approaches had similar performance over a 6-month interval from sampling to disease compared with the primary 12-month interval (appendix p 34). Using a 2% prior probability, PPVs approximately doubled without changing the overall pattern of results (appendix p 35).

Decision curve analysis of these testing approaches varied by setting (figure 2B). In high-burden settings, IGRA offered minimal additional net benefit over a treat-all strategy. *BATF2* offered greater net benefit than IGRA and, using prior probability of 1%, had the highest net benefit across threshold probabilities of between 0·4% and 2·2%, equating to an NWT to prevent a single case of 45–250. Above this threshold probability, in which NWT

	Cases	Controls	Total	AUROC (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	p value
Roe3	189	4982	5171	0·77 (0·73–0·81)	0·74 (0·67–0·79)	0·70 (0·69–0·71)	Reference
BATF2	189	4982	5171	0·77 (0·73–0·81)	0·74 (0·67–0·80)	0·69 (0·67–0·70)	0·72
FCGR1A/B	189	4981	5170	0·77 (0·73–0·81)	0·65 (0·58–0·72)	0·79 (0·78–0·80)	0·87
ANKRD22	138	3243	3381	0·77 (0·72–0·81)	0·57 (0·48–0·65)	0·86 (0·85–0·87)	0·88
GBP2	189	4982	5171	0·75 (0·71–0·79)	0·74 (0·67–0·79)	0·65 (0·64–0·66)	0·064
SERPING1	189	4979	5168	0·75 (0·71–0·79)	0·66 (0·59–0·73)	0·75 (0·73–0·76)	0·069

Performance metrics of single-gene transcripts with equivalent performance to the best multi-gene signature (Roe3) to discriminate individuals with subclinical tuberculosis from individuals who remained disease free at an interval of 12 months from sampling to disease. Equivalence to Roe3 was defined as an adjusted p value greater than 0·05 in the pairwise Delong test. Sensitivity and specificity are the maximum Youden index calculated from the one-stage meta-analysis. AUROC=area under the receiver operating curve.

Table 2: **Performance metrics of equivalent RNA biomarker single-gene transcripts for subclinical tuberculosis**

Figure 1: ***BATF2*** **expression and diagnostic accuracy for different disease states**

Expression (Z score transformed) of the best-performing single-gene transcript, *BATF2*, is shown as a combined boxplot and scatterplot, stratified by disease state. Disease state is classified according to International Consensus for Early Tuberculosis consensus definitions and descriptive terms are also shown. Boxes represent IQR and median values. Coloured dots represent individual samples and darker dots represent outliers. Area under the receiver operating characteristic curve with 95% CIs of *BATF2* to discriminate each disease state from individuals who remained disease free is also shown.

was less than 45, treating none was best. The combined approach (in which only those positive for both tests are offered preventive treatment) had only slightly greater net benefit than *BATF2* alone at threshold probabilities of 1–3%, equating to an NWT of 33–100. By contrast, in low-burden settings, IGRA outperformed *BATF2* across all threshold probabilities. IGRA offered the highest net benefit at threshold probabilities under 2%, equating to an NWT of over 50, whereas the combined approach offered the highest net benefit at threshold probabilities of 2–7%, equating to an NWT of 14–50. Using a prior probability of 2%, findings were similar, but net benefit for all strategies was shifted to the right on the threshold probability scale (appendix p 35). For example, in high-burden settings,

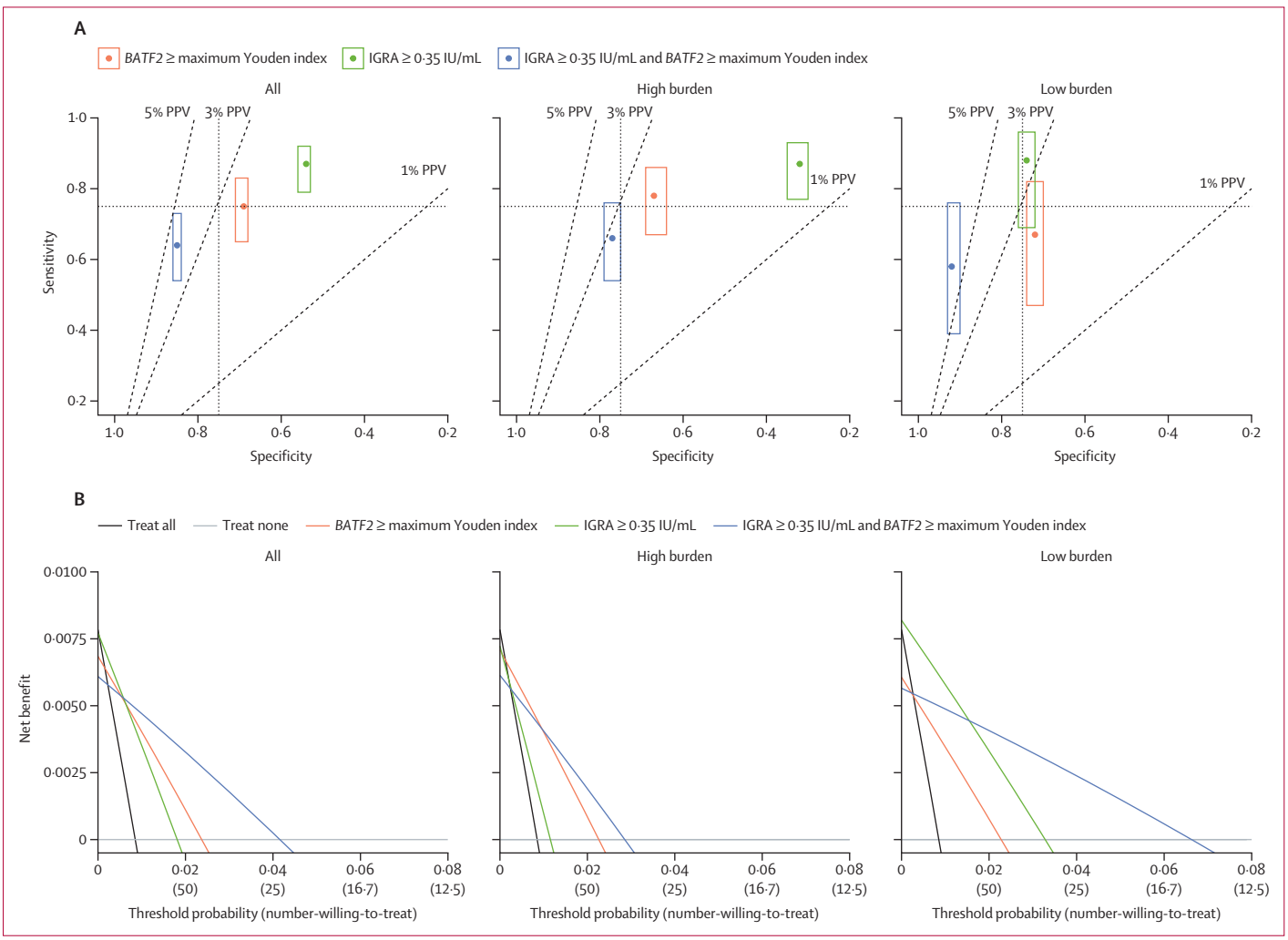

***Figure 2*: Diagnostic performance for subclinical tuberculosis of *BATF2*, IGRA, and a combined approach shown in receiver operating space and in a decision curve analysis, by setting**

(A) Diagnostic performance for subclinical tuberculosis over a 12-month interval of *BATF2* (threshold set at maximum Youden index), IGRA (QuantiFERON-TB Gold-in-tube or QuantiFERON-TB Gold Plus with threshold set at standard cutoff of 0·35 IU/mL), and a combined approach of *BATF2* and IGRA in the receiver operating space with sensitivity on the y-axis and 1–specificity on the x-axis, stratified by setting. Only participants with results for both tests were included in this analysis. Point estimates are shown (dots) with 95% CIs (boxes). Dotted lines represent the WHO minimum target product profile of 75% sensitivity and specificity for a tuberculosis progression test. Dashed lines represent PPVs of 1%, 3%, and 5%, based on a 1% prior probability. Underlying data are shown in the appendix (p 33). (B) Decision curve analysis in which each test is compared with default strategies of treating all or treating no people, stratified by setting. Threshold probability is the risk of tuberculosis disease at which a clinician or patient would opt for preventive treatment and is the reciprocal of the number-willing-to-treat to prevent a single case. Net benefit is calculated at a range of threshold probabilities as the true positive rate minus a weighted false positive rate, in which the weighting is the threshold probability. Since the contributing datasets included case–control analyses, the cumulative tuberculosis risk was fixed at 1%. IGRA=interferon-γ release assay. PPV=positive predictive value.

BATF2 offered greater net benefit over IGRA and a treat-all strategy at threshold probabilities of 0·7–4·5%, equating to an NWT of 22–143.

NNT estimates to prevent a single tuberculosis case using the testing approaches to stratify treatment, compared with a treat-all strategy, are shown in figure 3A. In high-burden settings, performing IGRA testing resulted in slightly lower NNT estimates than treating all (98 [95% CI 88–114] *vs* 125). Compared with IGRA, NNTs were significantly lower using *BATF2* (54 [95% CI 46–68]) or a combined approach (44 [34–59]). In low-burden settings, NNTs of IGRA and *BATF2* were similar (38 [32–51] *vs* 53 [40–80]). Using the combined approach resulted in a lower NNT (18 [13–33]) compared with *BATF2* alone. Due to superior specificity, lower NNT estimates were reached by IGRA (either alone or in combination with *BATF2*) in low-burden settings versus high-burden settings. Using a 2% prior probability roughly halved the NNT with a similar pattern (figure 3B).

In sensitivity analyses, a two-stage meta-analysis to calculate pooled AUROCs resulted in similar findings to the primary analysis (appendix p 36). Similarly, we observed similar AUROCs when including recipients of tuberculosis preventive therapy (excluding the participants in CORTIS-01 positive for Darboe11 who were randomly assigned to receive tuberculosis preventive therapy), including only one RNA sample per individual, or when including datasets from which the signatures were originally derived in the accuracy calculation for that signature (appendix p 36).

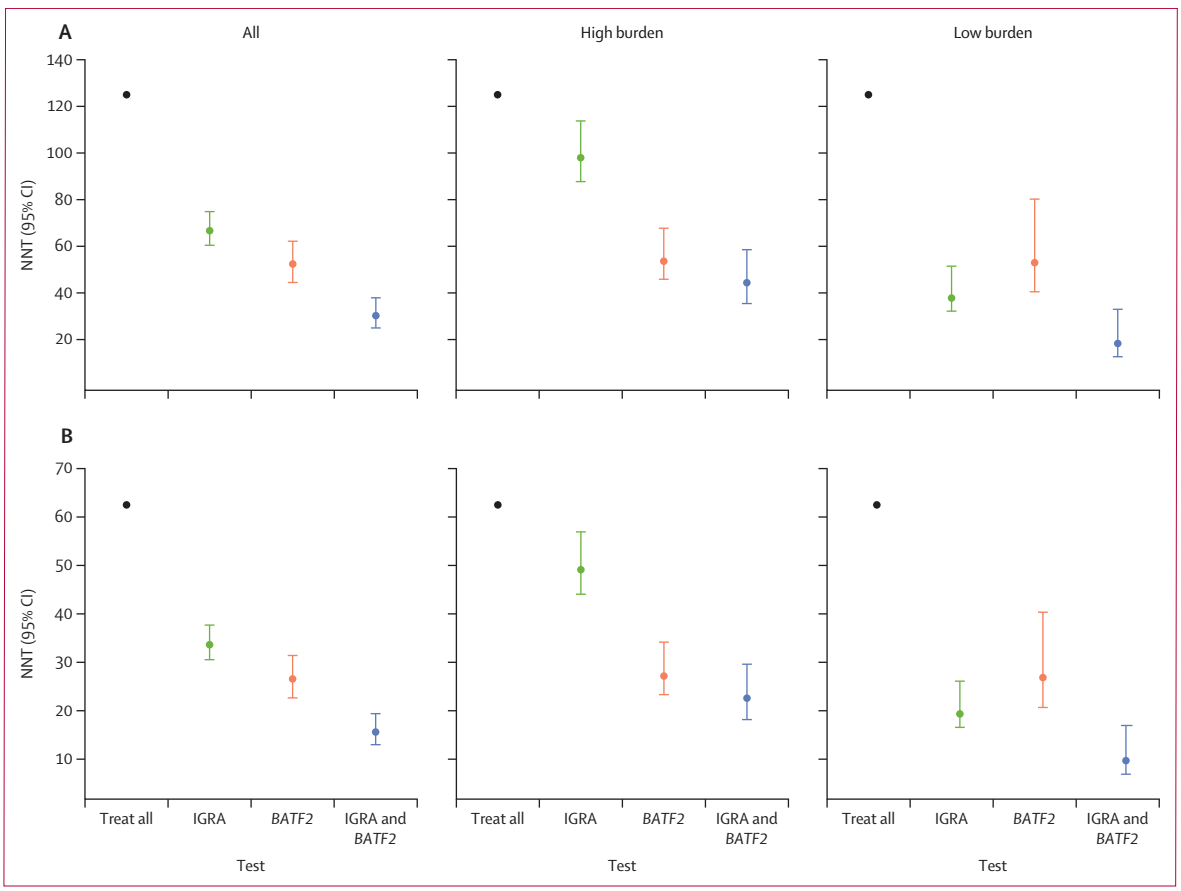

Figure 3: **NNT to prevent a single tuberculosis case using different testing strategies, by setting**
Estimated NNT with preventive treatment to prevent a single tuberculosis case is shown for the default strategy of treating all compared with test stratified treatment using *BATF2* (threshold set at maximum Youden index), IGRA (threshold set at standard cutoff of 0·35 IU/mL), and a combined approach of *BATF2* and IGRA. Point estimates are shown (dots) with 95% CIs (bars). An estimated preventive treatment effect of 80% was used. Since the contributing datasets included case–control analyses, the cumulative tuberculosis risk (prior probability) was fixed at 1% (A) and 2% (B). IGRA=interferon-γ release assay. NNT=number-needed-to-treat.

Discussion

To our knowledge, we report the largest pooled RNA biomarker analysis in subclinical tuberculosis to date, including over 6500 RNA samples, and the first comprehensive head-to-head analysis comparing single-gene transcripts with multi-gene signatures. This study is also the first evaluation of RNA biomarkers to align with the ICE-TB classification. We showed that five single-gene transcripts perform equivalently to the best performing multi-gene signature for differentiation of individuals with subclinical tuberculosis from those who remained disease free over 12 months. RNA biomarker performance was consistent across settings. By contrast, IGRA performance varied markedly, with poor performance in high-burden settings. In decision curve analysis, we show that, in high-burden settings, stratifying preventive treatment using RNA biomarker testing offers higher net benefit than using IGRA, which has minimal benefit over treating all. In low-burden settings, IGRA was the best single test to stratify treatment and approximated the WHO TPP, with greater net benefit than RNA biomarker testing. However, for tuberculosis programmes aiming to treat fewer than 50 people to prevent a tuberculosis case, a two-step combined testing approach improves specificity and is superior.

Development of the Cepheid MTB-HR prototype has shown that translation to a near-patient platform is feasible. Although the cost of such platforms is unknown, it is unlikely to exceed the WHO TPP maximum of US$100 per test, based on the cost of an IGRA.[33] Our findings could facilitate translation of RNA biomarker technology to clinical practice by encouraging commercial competition using measurement of any one of the best-performing single-genes at lower cost compared with measurement of multigene signatures. Nonetheless, none of the transcripts met the WHO TPP minimum sensitivity and specificity, even over a 12-month interval, although this target was almost achieved by IGRA in the low-burden setting (88% sensitivity and 74% specificity). In the high-burden setting, the WHO TPP seem an unrealistic aim that is unlikely to be achieved over 2 years with a transcriptional biomarker targeting early disease, although serial testing could improve overall performance. Combining different modalities of

tests might improve specificity, albeit at greater cost. A universal testing strategy might be challenging to attain if test performance is heterogeneous across settings; rather, tailored strategies might be required based on tuberculosis burden. In low-burden settings IGRA remains a useful test; an approach combining IGRA and RNA biomarkers shows additional promise and warrants further evaluation. However, in high-burden settings, the high prevalence of *M tuberculosis* sensitisation means that IGRA has poor specificity and thus minimal utility. Greater specificity might be achieved with better measures of recent (eg, less than 6 months) *M tuberculosis* infection, for example *M tuberculosis*-specific T-cell activation.[39] Alternatively, combining molecular approaches with radiological testing, such as digital chest radiographs,[40] as a method to detect macroscopic pathology might also improve performance. Moreover, as preventive treatment regimens become shorter and more acceptable, the NWT will likely increase, making a treat-all approach more attractive, particularly among individuals at high risk of progression to disease. Future studies are required to explore the trade-offs between tuberculosis risk and treatment acceptance, using current and truncated regimens, and to explore the clinical and cost-effectiveness of biomarker-stratified versus universal treatment approaches. Further trials are also required to identify optimal treatments for subclinical diseases states.

Our findings could provide some insights into host immune responses in early tuberculosis. The co-correlation of best-performing single-gene transcripts is consistent with previous findings of shared upstream interferon (IFN) and tumour necrosis factor (TNF) signalling pathways,[25] which explains why a single transcript is a sufficient measure of this immune response and combining these transcripts into multi-gene signatures does not offer orthogonal value. The consistent performance of RNA biomarkers across settings suggests that this is a common host response across populations. Likewise, the similar performance in CORTIS-HR suggests that this pathway is preserved in people living with HIV, although previous data have shown probable upregulation of common type-I IFN responses in people with untreated HIV.[41] Similarly, through IFN and TNF signalling, respiratory viral infections have been shown to increase tuberculosis signature scores,[42] which might partly account for the imperfect specificity of RNA biomarkers.[43] The imperfect sensitivity and the fall in discriminative performance after 12 months might be reflective of de-novo infection following signature measurement in high-burden settings, or that there is a minimal host response in earlier subclinical disease. The equivalent performance of multiple co-correlated RNA biomarkers, which has also been reported previously,[25,29,41] suggests that future discovery and validation of signatures using similar approaches is unlikely to yield a test with better performance. To date, discovery approaches have largely focused on identifying differentially expressed transcripts in tuberculosis, before combining these to form a discriminating multi-gene signature. Simplifying these to single-gene biomarkers has the added benefit of facilitating their integration into panels of blood RNA biomarkers for multinomial classification[44] that might overcome the specificity limitations of the binomial approach, for example by combining with the best-performing single-gene biomarkers of viral infections.[45]

An important strength of our analysis is that we adopted ICE-TB terminology for subclinical tuberculosis to facilitate comparisons across studies and biomarker domains. We used 6530 samples including 283 samples from subclinical tuberculosis cases, making it the largest analysis of RNA biomarkers for subclinical tuberculosis to date. Our data processing pipeline ensured batch correction within studies and integrated RNA sequencing and qPCR data into a pooled dataset, although our analyses were restricted to the pool of transcripts measured in at least one qPCR study. We also performed the first decision curve analysis, to our knowledge, to quantify the clinical utility of RNA biomarkers for subclinical tuberculosis and compare with existing tests. We also stratified our decision curve analysis by tuberculosis burden, which allowed granular assessment of clinical utility by setting. We performed multiple sensitivity analyses, including a two-stage IPD meta-analysis, to ensure our primary findings were robust.

A limitation of our study is that, although we included cohorts from high-burden and low-burden settings, there were few contributing countries (South Africa, UK, Brazil, The Gambia, and Ethiopia), with no representation from Asia, although large proportions of the UK studies were individuals of south Asian ethnicity. There were low numbers of subclinical cases over a 12-month interval in some studies, with both Leicester Contacts and CORTIS-HR reporting five cases each, which reflects the reality that tuberculosis is a rare outcome in longitudinal cohort studies. There were variations in case definitions between studies, which might have resulted in misclassification; however, this variation is reflective of real-world variations in clinical practice according to resource availability. As sputum sampling was not widely performed at baseline, there is also a risk of misclassification between subclinical, non-infectious and subclinical, infectious disease states; however, combining these groups in our primary analysis mitigated this risk. Furthermore, with the exception of CORTIS-01 and CORTIS-HR, evaluation of tuberculosis disease during follow-up was symptom-triggered, so additional cases of subclinical tuberculosis might have been missed, leading to underestimations of specificity for subclinical tuberculosis. We also acknowledge that there is no gold standard for the subclinical, non-infectious state and high-resolution investigations for macroscopic pathology, such as PET-CT,[46] were not performed. We therefore assumed that participants who developed tuberculosis within 12 months would have had macroscopic pathology at baseline, had high-resolution investigation been performed. However ongoing *M tuberculosis* exposure during follow-up, particularly in

high-transmission settings, might mean that disease cases within the primary 12-month interval could be attributable to new infection, leading to underestimated sensitivity for subclinical tuberculosis. Since subclinical tuberculosis can regress or undulate without treatment,[47] we could have underestimated specificity for subclinical, non-infectious tuberculosis that did not progress to clinical disease within 12 months. Future studies will be required to further evaluate the accuracy of candidate biomarkers for the subclinical, non-infectious state, once a scalable and widely accepted reference standard is established.

In summary, we have shown that several single-gene transcripts perform equivalently to multi-gene signatures to detect subclinical tuberculosis, which could simplify assays and encourage commercial competition. RNA biomarker performance is consistent across settings and exceeds performance of IGRA in high-burden settings but falls short of WHO benchmarks. A combination strategy with IGRA shows promise to enable more targeted preventive treatment in low-incidence settings.

Contributors

JG-B, MN, and RKG conceived the study. JG-B and RKG wrote the analysis plan with input from all authors. JG-B performed the analyses and wrote the first draft of the manuscript, supported by RKG and MN. JG-B and RKG verified the data. All authors had full access to the data. All authors contributed to the methods and interpretation. All authors had final decision to submit for publication and approved the final submitted manuscript.

Declaration of interests

MN holds a patent in relation to blood transcriptomic biomarkers of tuberculosis. TJS is co-inventor of patents of the Penn-Nicholson6 and Suliman4 signatures. All other authors declare no competing interests.

Data sharing

All data analysed in this study are publicly available through the original contributing studies.

Acknowledgments

We thank the study teams and participants from all of the primary studies included in this analysis (Adolescent Cohort Study, Correlate of Risk Targeted Intervention Study [CORTIS]-01, CORTIS-HR, Grand Challenges 6–74, Leicester Contacts, London Contacts, and REPORT-Brazil). RKG is funded by National Institute for Health Research (NIHR; NIHR303184), the British Medical Association Foundation for Medical Research, and by NIHR Biomedical Research Funding to UCL and UCLH. MN was supported by the Wellcome Trust (207511/Z/17/Z, 306550/Z/23/Z) and by NIHR Biomedical Research Funding to UCL and UCLH.

References

1 WHO. Global tuberculosis report 2023. https://www.who.int/publications-detail-redirect/9789240083851 (accessed May 28, 2024).
2 Litvinjenko S, Magwood O, Wu S, Wei X. Burden of tuberculosis among vulnerable populations worldwide: an overview of systematic reviews. *Lancet Infect Dis* 2023; **23:** 1395–407.
3 Coussens AK, Zaidi SMA, Allwood BW, et al. Classification of early tuberculosis states to guide research for improved care and prevention: an international Delphi consensus exercise. *Lancet Respir Med* 2024; **12:** 484–98.
4 Paton NI, Cousins C, Suresh C, et al. Treatment strategy for rifampin-susceptible tuberculosis. *N Engl J Med* 2023; **388:** 873–87.
5 Mahomed H, Hawkridge T, Verver S, et al. The tuberculin skin test versus QuantiFERON TB Gold in predicting tuberculosis disease in an adolescent cohort study in South Africa. *PLoS One* 2011; **6:** e17984.
6 Kik SV, Franken WPJ, Mensen M, et al. Predictive value for progression to tuberculosis by IGRA and TST in immigrant contacts. *Eur Respir J* 2010; **35:** 1346–53.
7 Zhou G, Luo Q, Luo S, et al. Interferon-γ release assays or tuberculin skin test for detection and management of latent tuberculosis infection: a systematic review and meta-analysis. *Lancet Infect Dis* 2020; **20:** 1457–69.
8 Melnychuk L, Perlman-Arrow S, Lisboa Bastos M, Menzies D. A systematic review and meta-analysis of tuberculous preventative therapy adverse events. *Clin Infect Dis* 2023; **77:** 287–94.
9 Berry MPR, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* 2010; **466:** 973–77.
10 Bloom CI, Graham CM, Berry MPR, et al. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. *PLoS One* 2013; **8:** e70630.
11 Kaforou M, Wright VJ, Oni T, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. *PLoS Med* 2013; **10:** e1001538.
12 Anderson ST, Kaforou M, Brent AJ, et al. Diagnosis of childhood tuberculosis and host RNA expression in Africa. *N Engl J Med* 2014; **370:** 1712–23.
13 Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* 2016; **4:** 213–24.
14 Roe JK, Thomas N, Gil E, et al. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. *JCI Insight* 2016; **1:** e87238.
15 Maertzdorf J, McEwen G, Weiner J 3rd, et al. Concise gene signature for point-of-care classification of tuberculosis. *EMBO Mol Med* 2016; **8:** 86–95.
16 Gjøen JE, Jenum S, Sivakumaran D, et al. Novel transcriptional signatures for sputum-independent diagnostics of tuberculosis in children. *Sci Rep* 2017; **7:** 5839.
17 Rajan JV, Semitala FC, Mehta T, et al. A novel, 5-transcript, whole-blood gene-expression signature for tuberculosis screening among people living with human immunodeficiency virus. *Clin Infect Dis* 2019; **69:** 77–83.
18 Kulkarni V, Queiroz ATL, Sangle S, et al. A two-gene signature for tuberculosis diagnosis in persons with advanced HIV. *Front Immunol* 2021; **12:** 631165.
19 Natarajan S, Ranganathan M, Hanna LE, Tripathy S. Transcriptional profiling and deriving a seven-gene signature that discriminates active and latent tuberculosis: an integrative bioinformatics approach. *Genes (Basel)* 2022; **13:** 616.
20 Chendi BH, Jooste T, Scriba TJ, et al. Utility of a three-gene transcriptomic signature in the diagnosis of tuberculosis in a low-endemic hospital setting. *Infect Dis (Lond)* 2023; **55:** 44–54.
21 Zak DE, Penn-Nicholson A, Scriba TJ, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet* 2016; **387:** 2312–22.
22 Suliman S, Thompson EG, Sutherland J, et al. Four-gene pan-African blood signature predicts progression to tuberculosis. *Am J Respir Crit Care Med* 2018; **197:** 1198–208.
23 Penn-Nicholson A, Mbandi SK, Thompson E, et al. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci Rep* 2020; **10:** 8629.
24 Roe J, Venturini C, Gupta RK, et al. Blood transcriptomic stratification of short-term risk in contacts of tuberculosis. *Clin Infect Dis* 2020; **70:** 731–37.
25 Gupta RK, Turner CT, Venturini C, et al. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. *Lancet Respir Med* 2020; **8:** 395–406.
26 Olbrich L, Verghese VP, Franckling-Smith Z, et al. Diagnostic accuracy of a three-gene *Mycobacterium tuberculosis* host response cartridge using fingerstick blood for childhood tuberculosis: a multicentre prospective study in low-income and middle-income countries. *Lancet Infect Dis* 2024; **24:** 140–49.

27 Gupta-Wright A, Ha H, Abdulgadar S, et al. Evaluation of the Xpert MTB Host Response assay for the triage of patients with presumed pulmonary tuberculosis: a prospective diagnostic accuracy study in Viet Nam, India, the Philippines, Uganda, and South Africa. *Lancet Glob Health* 2024; **12:** e226–34.

28 Scriba TJ, Fiore-Gartland A, Penn-Nicholson A, et al. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infect Dis* 2021; **21:** 354–65.

29 Mendelsohn SC, Fiore-Gartland A, Penn-Nicholson A, et al. Validation of a host blood transcriptomic biomarker for pulmonary tuberculosis in people living with HIV: a prospective diagnostic and prognostic accuracy study. *Lancet Glob Health* 2021; **9:** e841–53.

30 Mendelsohn SC, Andrade BB, Mbandi SK, et al. Transcriptomic signatures of progression to tuberculosis disease among close contacts in Brazil. *J Infect Dis* 2024; **230:** e1355–65.

31 Mendelsohn SC, Mbandi SK, Fiore-Gartland A, et al. Prospective multicentre head-to-head validation of host blood transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR. *Commun Med (Lond)* 2022; **2:** 26.

32 Gupta RK, Calderwood CJ, Yavlinsky A, et al. Discovery and validation of a personalized risk predictor for incident tuberculosis in low transmission settings. *Nat Med* 2020; **26:** 1941–49.

33 WHO. 3.0 IGO consensus meeting report: development of a target product profile (TPP) and a framework for evaluation for a test for predicting progression from tuberculosis infection to active disease. World Health Organization, 2017.

34 Martinez L, Cords O, Horsburgh CR, Andrews JR, Pediatric TB Contact Studies Consortium. The risk of tuberculosis in children after close exposure: a systematic review and individual-participant meta-analysis. *Lancet* 2020; **395:** 973–84.

35 Du Y, He Y, Zhang H, et al. Declining incidence rate of tuberculosis among close contacts in five years post-exposure: a systematic review and meta-analysis. *BMC Infect Dis* 2023; **23:** 373.

36 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006; **26:** 565–74.

37 Viechtbauer W. Conducting meta-analyses in R with the metafor Package. *J Stat Softw* 2010; **36:** 1–48.

38 Singhania A, Verma R, Graham CM, et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat Commun* 2018; **9:** 2308.

39 Mpande CAM, Musvosvi M, Rozot V, et al. Antigen-specific T-cell activation distinguishes between recent and remote tuberculosis infection. *Am J Respir Crit Care Med* 2021; **203:** 1556–65.

40 Tan Q, Huang CC, Becerra MC, et al. Chest radiograph screening for detecting subclinical tuberculosis in asymptomatic household contacts, Peru. *Emerg Infect Dis* 2024; **30:** 1115–24.

41 Mann T, Gupta RK, Reeve BWP, et al. Blood RNA biomarkers for tuberculosis screening in people living with HIV before antiretroviral therapy initiation: a diagnostic accuracy study. *Lancet Glob Health* 2024; **12:** e783–92.

42 Mulenga H, Musvosvi M, Mendelsohn SC, et al. Longitudinal dynamics of a blood transcriptomic signature of tuberculosis. *Am J Respir Crit Care Med* 2021; **204:** 1463–72.

43 Noursadeghi M, Gupta RK. New insights into the limitations of host transcriptional biomarkers of tuberculosis. *Am J Respir Crit Care Med* 2021; **204:** 1363–65.

44 Habgood-Coote D, Wilson C, Shimizu C, et al. Diagnosis of childhood febrile illness using a multi-class blood RNA molecular signature. *Med (N Y)* 2023; **4:** 635–54.e5.

45 Gupta RK, Rosenheim J, Bell LC, et al. Blood transcriptional biomarkers of acute viral infection for detection of pre-symptomatic SARS-CoV-2 infection: a nested, case-control diagnostic accuracy study. *Lancet Microbe* 2021; **2:** e508–17.

46 Esmail H, Lai RP, Lesosky M, et al. Characterization of progressive HIV-associated tuberculosis using 2-deoxy-2-[^{18}F]fluoro-D-glucose positron emission and computed tomography. *Nat Med* 2016; **22:** 1090–93.

47 Richards AS, Sossen B, Emery JC, et al. Quantifying progression and regression across the spectrum of pulmonary tuberculosis: a data synthesis study. *Lancet Glob Health* 2023; **11:** e684–92.