Prospective validation of host transcriptomic biomarkers for pulmonary tuberculosis by real-time PCR

Simon C. Mendelsohn, MBChB¹, Stanley Kimbung Mbandi, PhD¹, Andrew Fiore-Gartland, PhD², Adam Penn-Nicholson, PhD¹, Munyaradzi Musvosvi, PhD¹, Humphrey Mulenga, MPH¹, Michelle Fisher, PhD¹, Katie Hadley, PhD¹, Mzwandile Erasmus, BSc¹, Onke Nombida, GT¹, Michèle Tameris, MBChB¹, Gerhard Walzl, MBChB³, Kogieleum Naidoo, PhD^{4,5}, Gavin Churchyard, PhD^{6,7,8}, Mark Hatherill, MD¹, Thomas J. Scriba, DPhil¹, The Cross-sectional TB Cohort Study Team*, and The CORTIS Study Team*

¹South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, and Division of Immunology, Department of Pathology, University of Cape Town, 7925, Cape Town, South Africa.

²Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA.

³DST/NRF Centre of Excellence for Biomedical TB Research; South African Medical Research Council Centre for TB Research; Division of Molecular Biology and Human Genetics, Department of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, 7505, Cape Town, South Africa.

⁴Centre for the AIDS Programme of Research in South Africa (CAPRISA), 4001, Durban, South Africa.

⁵MRC-CAPRISA HIV-TB Pathogenesis and Treatment Research Unit, Doris Duke Medical Research Institute, University of KwaZulu-Natal, 4001, Durban, South Africa.

⁶The Aurum Institute, 2194, Johannesburg, South Africa.

⁷School of Public Health, University of Witwatersrand, 2193, Johannesburg, South Africa.

⁸Department of Medicine, Vanderbilt University, Nashville, TN 37232, USA.

* See supplement

Word count: 6,821 words (48,585 characters including spaces)

Running title: Blood mRNA biomarkers for pulmonary TB

Corresponding author: Thomas J Scriba

Email: Thomas.Scriba@uct.ac.za **Telephone:** +27 (0) 21 406 6427

Postal address: South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine, Division of Immunology, Department of Pathology, University of Cape Town, Anzio Road, Observatory, 7925, Cape Town, South Africa

Abstract

We tested performance of host-blood transcriptomic tuberculosis (TB) signatures for active case-finding. Among 20,207 HIV-uninfected and 963 HIV-infected adults screened; 2,923 and 861 were enrolled from five South African communities. Eight signatures were measured by microfluidic RT-qPCR and participants were microbiologically-investigated for pulmonary TB at baseline, and actively surveilled for incident disease through 15 months. Diagnostic AUCs for 61 HIV-uninfected (weighted-prevalence 1.1%) and 10 HIV-infected (prevalence 1.2%) prevalent TB cases for the 8 signatures were 0.63–0.79 and 0.65–0.88, respectively. Thereafter, 24 HIV-uninfected and 9 HIV-infected participants progressed to incident TB (1.1 and 1.0 per 100 person-years, respectively). Prognostic AUCs through 15-months follow-up were 0.49–0.66 and 0.54–0.81, respectively. Prognostic performance for incident TB occurring within 6-12 months in HIV-negative participants was higher for all signatures. None of the signatures met WHO Target Product Profile criteria for a triage test to diagnose asymptomatic TB; most signatures met the criteria for symptomatic TB. Prognostic accuracy of most signatures for incident TB within six months of testing met the criteria for an incipient TB test.

Keywords

Diagnostic / prognostic/ transcriptional / signature / Tuberculosis

Introduction

An estimated 2.9 million people with tuberculosis (TB) disease went undiagnosed or unreported in 2019¹, partly because passive case-finding strategies rely primarily on self-presentation of people for healthcare when they experience TB symptoms, and dependence on sputum for microbiological testing. TB prevalence surveys in Africa and Asia demonstrate that approximately 50% of TB cases are asymptomatic, highlighting a major blind-spot in the symptom-dependent case-finding strategies, which may perpetuate significant *Mycobacterium tuberculosis* (*Mtb*) transmission.²⁻⁵ Currently available triage tests and passive case-finding strategies for symptomatic disease are thus inadequate to find these missing cases among people who do seek healthcare. Effective active case-finding tools for screening of asymptomatic individuals in the community are urgently needed for earlier detection, further microbiological testing, and treatment, in order to interrupt the spread of *Mtb*.

Computer-aided detection for automated interpretation of digital chest radiographs⁶ is one such approach for rapid and cheap mass screening, particularly in resource-constrained settings. Mobile sputum Xpert MTB/RIF screening also performs well as a community-based case-finding strategy, but requires significant resources and is dependent on availability of a sputum sample.⁷ Tests for detection of *Mtb*-exposed individuals at high risk of progression to TB disease, particularly in high burden settings, are also urgently needed to guide TB preventive therapy (TPT). Interferon-γ release assays (IGRA) and tuberculin skin testing lack specificity in endemic settings due to high rates of prior *Mtb* exposure, which result in both overtreatment of those unlikely to progress to disease and missed “incipient” cases.⁸ IGRA are also laborious, and require laboratory facilities and skilled technicians.

We previously developed a predictive correlate of risk signature, RISK11, for diagnosis of TB disease and for identification of individuals at high risk of progression to active TB disease.^{9,10} Diagnostic and prognostic performance of RISK11 as a PCR-based screening test was prospectively validated in HIV-uninfected adults¹¹ and people living with HIV (PLHIV)¹². This signature is not ideal for further development due to its large size (48 primer-probe assays). Fortunately, several refined, parsimonious diagnostic and prognostic transcriptomic signatures, which are more amenable to translation to point-of-care testing, have been developed.¹³⁻¹⁶ However, most transcriptomic signature discovery and validation studies use a case-control design with stringent exclusion criteria, resulting in spectrum bias and best-case discriminative performance, which is not reproducible in prospective studies in heterogeneous real-world settings. Transcriptomic signatures are predominantly measured by microarray or RNA-sequencing, techniques which are not suitable for use at the point of care. There are also few studies which have performed unbiased head-to-head comparisons of signature performance¹³⁻¹⁵ and none for

active case-finding in a community setting. It is therefore not clear which, if any, signatures offer potential in this context and should be advanced along the developmental pipeline.

We recently showed that HIV and upper respiratory tract viruses result in an increase in RISK11 score and hypothesised that this was due to the upregulation of interferon-stimulated genes (ISG) as part of an inflammatory response typically induced in TB and viral infections.^{10,12,17} Further, RISK11 was able to differentiate individuals with TB disease from those with bacterial or no upper respiratory tract microorganisms, but was not able to discriminate between those with TB from those with upper respiratory tract viruses. Most TB transcriptomic signatures primarily detect elevated ISG expression. The effect of HIV and respiratory viruses on scores and performance of these other signatures, including those which do not encompass ISGs such as the Thompson5 (RESPONSE5) signature¹⁸, has not been adequately explored.

In this paper we report the results of a multi-centre observational study, which prospectively tested the performance of eight parsimonious host-blood TB transcriptomic signatures measured by microfluidic real-time (RT) qPCR, for diagnosing prevalent pulmonary TB and predicting progression to incident TB disease, in a predominantly asymptomatic community cohort of HIV-uninfected and HIV-infected South African adults. As secondary aims, we compared signature diagnostic performance in symptomatic and asymptomatic participants, and benchmarked parsimonious signature accuracy against RISK11 and the World Health Organization (WHO) Target Product Profile (TPP) criteria for triage and prognostic TB tests.^{19,20} As an exploratory aim, we also explored the effect of HIV and upper respiratory tract microorganisms on transcriptomic signature scores.

Methods

Study design and participants

In this diagnostic and prognostic accuracy study, we evaluated parsimonious blood-based host transcriptomic signature performance in four South African cohorts (**Figure S1**). We first translated the signatures to a RT-qPCR platform and validated performance in a case-control study of HIV-uninfected and HIV-infected individuals with and without pulmonary TB, the Cross-sectional TB Cohort (CTBC) study (**Figure S1A**). We then prospectively tested signature performance in HIV-uninfected volunteers enrolled in the CORTIS-01 trial (ClinicalTrials.gov: NCT02735590; **Figure S1B**) and PLHIV enrolled in the observational CORTIS-HR study (**Figure S1D**). Finally, we evaluated the effects of upper respiratory microorganisms on transcriptomic signatures in a subset of HIV-uninfected participants screened for the CORTIS-01 trial (**Figure S1C**). This study is reported in accordance with the Standards for Reporting Diagnostic Accuracy Studies initiative guidelines²¹ and the study protocol was approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (HREC 812/2017). Written informed consent was obtained from all participants and all experiments conformed to the principles set out in the WMA Declaration of Helsinki and the US Department of Health and Human Services Belmont Report.

Cross-sectional TB Cohort

The CTBC case-control study was previously described.²² Briefly, adults aged ≥ 18 years with newly diagnosed pulmonary TB confirmed by sputum Xpert MTB/RIF (Cepheid, Sunnyvale, CA, USA) and/or liquid mycobacterial culture (Mycobacteria Growth Indicator Tube [MGIT], BACTEC, Beckton Dickinson, Franklin Lakes, NJ, USA) were enrolled at primary healthcare clinics in Worcester and Masiphumelele, South Africa. Healthy, asymptomatic community controls were recruited from the same areas. HIV-infection was diagnosed with the Determine HIV1/2 test (Alere, Waltham, MA, USA). All participants provided written, informed consent and the protocols were approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (HREC 126/2006 and 288/2008).

Correlate of Risk Targeted Intervention Study (CORTIS)

The CORTIS-01 clinical trial in HIV-uninfected individuals and CORTIS-HR observational study in PLHIV have previously been described.^{11,12} Briefly, healthy adult volunteers without clinical suspicion of TB, residing in five TB endemic communities in South Africa (Durban, Klerksdorp, Ravensmead, Rustenburg, and Worcester), were recruited through word-of-mouth, house-to-house visits, and liaison with non-governmental organisations. Recruitment did not target symptomatic individuals seeking health care or other high-risk groups. Eligible participants aged 18–59 years were without comorbidities (except for HIV) and did not have known TB disease, or

household exposure to individuals with multi-drug resistant TB, within the prior three years. The RISK11 signature was measured at enrolment, and a pre-specified score threshold of 60% was used to differentiate RISK11-positive from RISK11-negative participants. The HIV-uninfected CORTIS-01 population was enriched for participants at high risk of TB disease through enrolment of all eligible RISK11-positive individuals, and randomisation of RISK11-negative participants to enrolment or non-enrolment at a 1:7-9 ratio. RISK11-positive participants were randomised to receive TPT (weekly high-dose isoniazid and rifapentine for three months, 3HP) or no intervention; RISK11-negative participants were randomised to no intervention or excluded from the study. In the CORTIS-HR observational study, PLHIV were enrolled irrespective of RISK11 status and referred for standard of care antiretroviral and isoniazid preventive therapy.

All CORTIS participants provided two spontaneously expectorated sputum samples at enrolment and at the end of study visits. End of study visits were performed at month 15 of follow-up or at an earlier timepoint for withdrawn participants. Both samples collected at enrolment from HIV-uninfected participants (CORTIS-01) were tested for *Mtb* using Xpert MTB/RIF, and from HIV-infected participants (CORTIS-HR) using Xpert MTB/RIF and MGIT culture. The two samples collected at the end of study visit were tested using Xpert MTB/RIF or Xpert Ultra (Cepheid), and MGIT culture. In addition, symptom-triggered TB investigations (two sputum samples; one for Xpert MTB/RIF and one for MGIT culture) were performed at six routine study visits through 15-months follow-up. TB symptoms included at least one of: persistent unexplained cough, night sweats, fever, or weight loss for 2 weeks or more, or any haemoptysis. Participants diagnosed with microbiologically-confirmed TB were withdrawn from the study and referred for curative treatment. All participants provided written informed consent and study protocols were approved by the institutional human research ethics committees of the five participating sites.

Respiratory pathobionts cohort

The respiratory pathobionts cohort, a subset of participants screened for eligibility for the CORTIS-01 trial at the Worcester site, has previously been described.¹⁷ Pathobionts encompass pathogenic and commensal microorganisms. Briefly, HIV-uninfected participants were consecutively enrolled in this sub-study irrespective of CORTIS-01 enrolment, or signs and symptoms of upper respiratory tract infections. Paired nasopharyngeal and oropharyngeal flocked swabs (FLOQSwabs, Copan Diagnostics, Murrieta, CA, USA) were collected and stored in Primestore buffer (Longhorn Vaccines and Diagnostics, San Antonio, TX, USA) at -80°C, and viral and bacterial nucleic acid was later extracted (Qiasymphony Virus/Bacteria Mini Kit, Qiagen, Hilden, Germany) and quantified using a multiplex RT-qPCR assay kit (Respiratory Pathogens 33 Kit, Fast Track Diagnostics, Luxembourg) on the CFX96 Touch System lightcycler platform (Bio-Rad, Hercules, CA, USA), according to the manufacturer's instructions. Participants co-enrolled into the CORTIS-01 trial were investigated for TB at baseline; those only enrolled into the respiratory pathobionts

cohort were not investigated for TB (**Figure S1C**). All participants provided written, informed consent and the protocols were approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (HREC 327/2017).

Blood sample collection and RNA extraction

Venous whole blood was collected in PAXgene tubes (PreAnalytiX, Hombrechtikon, Switzerland) at enrolment in all cohorts, frozen at -20°C, and shipped to the South African Tuberculosis Vaccine Initiative (SATVI) Cape Town laboratory. For the CTBC study, RNA was manually extracted with the PAXgene blood RNA kit (Qiagen) according to the manufacturer's instructions, stored at -80 °C, and later used for transcriptomic analysis. For the other cohorts, RNA was extracted using a high-throughput, standardised, and reproducible fully automated procedure on the Freedom EVO 150 robotic platform (Tecan, Männedorf, Switzerland) with the Maxwell SimplyRNA kit (Promega, Madison, WI, USA). One aliquot of RNA was used immediately for cDNA synthesis and measurement of the RISK11 signature, and a second aliquot was stored at -80 °C, and later used to measure the panel of parsimonious transcriptomic signatures.

Measurement of TB transcriptomic signatures

Eight parsimonious diagnostic, prognostic, and treatment response host-blood TB transcriptomic signatures were pragmatically selected for inclusion in this head-to-head validation alongside RISK11 based on availability of validated performance data, number of transcripts (six or fewer), and availability of target sequences for custom primer-probe design or predesigned TaqMan assay. As a naming convention, we use the first author and number of genes included in the signature: Francisco2²³, Maertzdorf4²⁴ (also known as DIAG4), Penn-Nicholson6 (also known as RISK6)²², Suliman4 (also known as RISK4)²⁵, Roe1²⁶ (BATF2 only), Roe3²⁷, Sweeney3²⁸, and Thompson5¹⁸ (also known as RESPONSE5). The Herberg2 signature²⁹, designed to discriminate between viral and bacterial infection in febrile children, was also included as a potential discriminator between viral and mycobacterial infection in the respiratory pathobionts cohort. Signatures were measured and scores calculated as previously reported (**Tables S1-S6**).

Following cDNA-synthesis with EpiScript reverse transcriptase (Lucigen, Middleton, WI, USA), genes of interest (**Tables S2-S3**) were pre-amplified using pools of TaqMan primer-probe assays (Thermo Fisher Scientific, Waltham, MA, USA), and gene expression (raw cycle threshold, Ct) quantified by microfluidic multiplex RT-qPCR using either Fluidigm (San Francisco, CA, USA) 96.96 (96 samples multiplexed with 96 primer-probe assays) or 192.24 (192 samples multiplexed with 24 primer-probe assays) Gene Expression chips on the BioMark HD instrument (Fluidigm). Further details regarding primer-probe assay selection and qualification, data quality control, and signature score measurement are provided in the **Supplemental Methods**.

Endpoints

For the CTBC study and respiratory pathobionts sub-study, TB disease was defined by a single sputum sample positive for *Mtb* on either Xpert MTB/RIF and/or liquid culture at enrolment. The coprimary endpoints in the CORTIS studies were baseline prevalent TB disease and incident disease through 15 months follow-up confirmed by a positive Xpert MTB/RIF, Xpert Ultra, or MGIT culture, on two or more separate sputum samples collected within any 30-day period. The secondary endpoint was microbiologically-confirmed TB disease on at least one sputum sample. All Xpert Ultra trace positive results were excluded from the analysis because of the risk of false positives.

Statistical analyses

The statistical analysis plan for the CORTIS sub-studies was pre-specified (**Appendix p20**) and all statistical analyses were done in R (Boston, MA, USA), version 3.6.1. Sample sizes for this sub-study were not pre-specified, but were based on accumulated enrolment of the CTBC and CORTIS parent studies. Performance of the signatures was evaluated on their ability to diagnose prevalent TB disease at baseline and predict progression to incident TB disease through 15-month follow up. For all analyses, HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) participants were considered independently. Diagnostic performance at baseline was assessed in all participants enrolled in the CORTIS studies with available PAXgene samples. Thereafter, participants meeting the primary prevalent TB endpoint definition, participants who did not attend further follow-up visits, and participants randomized to the RISK11-positive 3HP group (CORTIS-01 only) were excluded from the primary endpoint prognostic performance analysis. Participants were followed for up to 15 months; those who discontinued follow-up prior to 15 months and did not meet the primary endpoint definition were censored at their final study visit or last negative sputum sample collection but included in the prognostic analysis.

Due to enrichment of RISK11-positive individuals in the CORTIS-01 enrolled population, analyses of signature performance required participant weighting to obtain estimates applicable to the screened population, effectively upweighting RISK11-negative participants in the CORTIS-01 analyses (i.e. inverse probability weighting). Enrolment in the CORTIS-HR study and respiratory pathobionts sub-study were independent of RISK11 status; analyses for these cohorts are not weighted.

For evaluation of diagnostic performance in the CTBC, CORTIS-HR, and respiratory pathobionts studies, area under the receiver operating characteristic curve (ROC) curve (AUC) was generated using the pROC package³⁰ in R. Signature performance (AUC) between HIV-infected and HIV-uninfected participants in the CTBC study was compared using methods described by DeLong et

al.³¹ Prognostic performance metrics through 15 months follow-up in CORTIS-HR were calculated by use of non-parametric methods for time-dependent ROC curve analysis from survival data using the R `survAM.estimate` function in the `survAccuracyMeasures` package³² in R. Binary weighted ROC analysis was performed in the CORTIS-01 cohort to evaluate both diagnostic and prognostic performance. In a post-hoc analysis, prognostic performance in CORTIS-01 was qualitatively compared between 6, 12, and 15 months of follow-up for the primary endpoint by use of identical methods.

There were no pre-defined score thresholds for the parsimonious signatures; optimal signature score thresholds were calculated using the minimum Euclidean distance to 100% sensitivity and specificity, i.e. the point closest to the top-left of the ROC curve, calculated for each cut-point (c) as follows: $\sqrt{[(1 - \text{Sensitivity}_{(c)})^2 + (1 - \text{Specificity}_{(c)})^2]}$.³³ Signature diagnostic and prognostic accuracy was also benchmarked against the minimal WHO TPP for a community-based TB triage or referral test¹⁹ (specificity 70% or sensitivity 90%) and the minimal WHO TPP for an incipient TB test, to predict progression to TB disease²⁰ (specificity 75% or sensitivity 75%), respectively.

Sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) at each threshold were calculated using binary endpoint indicators and standard formulae. Estimates of PPV and NPV were computed based on the observed incidence of TB in the enrolled study population. The 95% confidence intervals (CIs) on diagnostic and prognostic performance estimates were calculated with a non-parametric percentile bootstrap with 10,000 resamples.³⁴ Bootstrap sampling was stratified by RISK11 status in the CORTIS-01 cohort, but not stratified in the CTBC, CORTIS-HR, or respiratory pathobionts cohorts.

Signature score distribution was described using median and interquartile range. Differences between groups were calculated using the Mann-Whitney U test and corrected for multiple comparisons by use of the Benjamini-Hochberg Procedure.³⁵ Spearman's rank coefficient was used to report correlation between signature scores; only samples with all assays passing were include in the correlation matrices; and Roe1 and Roe3 signatures were multiplied by -1 to obtain a positive correlation for all signatures. An alpha of <0.05 was considered significant in all analyses.

Blinding

Transcriptomic signature scores were measured by laboratory personnel who were blinded to participant TB status. Participants and study staff responsible for TB screening were blinded to transcriptomic signature scores. The statistical analysis team had access to clinical and demographic data (including TB status), but were blinded to signature scores, to allow data cleaning and preparation of analysis scripts prior to database lock. Signature scores and TB

305 microbiology results were maintained in different files, which were only integrated after the study
 306 database had been cleaned and locked, and group allocations unblinded in CORTIS-01.

307

308 *Role of the funding sources*

309 The funders of this sub-study had no role in protocol development, study design, data collection,
 310 data analysis, data interpretation, or writing of the report.

Results

Transcriptomic signatures discriminate symptomatic pulmonary TB patients from healthy, asymptomatic controls

We translated measurement of eight parsimonious TB transcriptomic signatures to a microfluidic RT- qPCR platform using TaqMan primer-probe assays for quantification of transcripts, using previously described methods (**Table S1-6**). All new primer-probe assays passed qualification with efficient linear amplification (**Table S3**). We then validated the performance of the signatures in the CTBC study, a case-control cohort of HIV-uninfected and HIV-infected individuals with or without pulmonary TB. Participants included 114 HIV-uninfected (53 TB cases and 61 asymptomatic controls) and 86 HIV-infected (45 TB cases and 41 asymptomatic controls) adults (**Figure S1A**). The performance of the Penn-Nicholson6 (RISK6) and RISK11 signatures in this cohort was previously reported²², but is included here for comparison. We found that performance of all signatures to discriminate between HIV-uninfected South African adults with TB from healthy controls was excellent, with AUCs ranging from 0.89 (Thompson5, 95% CI 0.82–0.95) to 0.97 (RISK11 and Suliman4, both with 95% CI 0.93–1) (**Figure S2**). All signatures met the minimal WHO triage test TPP benchmark criteria in this group (**Table S7A**). Discriminative performance was generally lower in PLHIV, with AUCs ranging from 0.78 (Francisco2, 95% CI 0.68–0.88) to 0.96 (Suliman4, 95% CI 0.90–1), and was significantly worse ($p<0.05$) for RISK11 ($p=0.027$), Francisco2 ($p=0.0095$), and Sweeney3 ($p=0.021$) (**Figure S2**). Only Penn-Nicholson6 and Suliman4 met the minimal WHO triage test TPP benchmark criteria in PLHIV (**Table S7B**). Having validated the signatures for microfluidic RT-qPCR in a case-control context, we next wished to prospectively compare signature performance in a real-world setting.

Recruitment of prospective community screening cohorts

In the CORTIS-01 trial,¹¹ 20,207 adults were screened at five geographically distinct community sites throughout South Africa; 16,248 HIV-uninfected individuals met inclusion for the CORTIS-01 study, of whom 2,923 participants were randomized and all were included in this parsimonious signature sub-study (**Figure S1B** and **Table S8**). A further 963 adults were screened and 861 PLHIV met inclusion for the CORTIS-HR study¹² and all were included in this sub-study (**Figure S1D** and **Table S9**). There were 61 HIV-uninfected (weighted prevalence 1.1%, 95% CI 0.8–1.6) and 10 HIV-infected (prevalence 1.2%, 95% CI 0.6–2.1) baseline primary endpoint TB cases; over 80% of prevalent TB cases in the HIV-uninfected cohort, and 70% in the HIV-infected cohort, were asymptomatic. Thereafter, 362 HIV-uninfected participants (350 in the 3HP arm, 10 lost to follow-up, and 2 withdrawals) and 5 PLHIV (4 lost to follow-up and 1 withdrawal) were excluded from analysis of prognostic performance for incident TB, such that 2,500 HIV-uninfected and 846 PLHIV were included. Through 15 months follow-up, 24 HIV-uninfected participants (weighted incidence 1.1 per 100 person-years, 95% CI 0.6–1.5) and 9 PLHIV (incidence 1.0 per 100 person-years, 95%

CI 0.3–1.6) progressed to incident TB. Participants who discontinued follow-up prior to 15 months were included in the prognostic assessment, but censored at their final study visit.

Measurement of parsimonious transcriptomic signatures in the prospective cohorts

The RISK11 signature assay panel (**Table S2**) was measured in all randomised HIV-uninfected CORTIS-01 trial participants (N=2,923) and in all enrolled HIV-infected CORTIS-HR participants who had baseline RNA samples available 857/861 (99.5%). RISK11 performance in these cohorts was previously reported^{11,12}, but is included here for comparison without a pre-specified score threshold. The parsimonious transcriptomic signature assay panel (**Table S3**) was measured in 2,904/2,923 (99.3%) HIV-uninfected and 858/861 (99.7%) HIV-infected participants who had additional baseline RNA aliquots available. The parsimonious signature failure rate ranged from 1.4% (RISK6) to 25.3% (Suliman4) in the HIV-uninfected cohort and from 5.2% (RISK6) to 32.1% (Suliman4) in PLHIV (**Tables S10-11**). Failed signature results for individual samples due to failed RT-qPCR reactions were not repeated, and were excluded from analysis (details in **Supplemental Results**).

Median signature scores for RISK11, Francisco2, Maertzdorf4, RISK6, Suliman4, Sweeney3, and Thompson5 at baseline were highest amongst HIV-uninfected participants with symptomatic (clinical) TB, and successively lower in participants with asymptomatic (sub-clinical) and incident TB (participants who progressed to TB disease), and symptomatic and asymptomatic participants without TB (**Figure 1**). The opposite was observed for the Roe1 and Roe3 signatures; this is related to data processing rather than real differences in signature biology, and a similar trend would be obtained by additively inverting the scores. A similar progression was observed for most signatures amongst PLHIV, although differences in the TB spectrum were not as marked, especially for Francisco2 and Sweeney3 (**Figure 2**).

Scores for all of the signatures were significantly correlated ($p > 0.05$; **Figure S3K and S4K**); Parsimonious signature Spearman-rank correlation with RISK11 ranged from 0.41–0.90 in HIV-uninfected and 0.25–0.88 in HIV-infected cohorts. Signatures with overlapping genes, such as (1) Roe1, Roe3, and RISK11, or (2) Francisco2 and Sweeney3, were highly correlated (ρ 0.84–0.96). The Thompson5 (RESPONSE5) treatment response monitoring signature had the poorest correlation with other signatures in HIV-uninfected (0.22–0.56) and HIV-infected (0.16–0.46) cohorts.

Similar diagnostic performance in HIV-uninfected and HIV-infected individuals, and inferior performance for detecting subclinical TB

Diagnostic performance for primary endpoint prevalent TB was moderate for all signatures, ranging from an AUC of 0.63 to 0.79 in the HIV-uninfected cohort (**Figure 1 and Table S10**). Primary

endpoint diagnostic performance estimates in PLHIV (AUC 0.65–0.88) were similar or exceeded performance in the HIV-uninfected cohort for most signatures (**Figure 2 and Table S11**), except the Thompson5 and Sweeney3 signatures, which appeared to perform poorly in PLHIV (lower 95% confidence bound crossing 0.50). When disaggregated by TB symptom status, diagnostic performance for detection of asymptomatic (subclinical) TB among all asymptomatic participants, (AUCs ranging from 0.61 to 0.78) was inferior to performance for detection of symptomatic (clinical) TB among all symptomatic participants (AUCs ranging from 0.85–0.98) for all signatures in HIV-uninfected participants (**Figure 1**). AUC estimates for detecting symptomatic TB were also higher for most signatures (RISK11, Maertzdorf4, Roe3, Suliman4, and Thompson5, with AUCs ranging from 0.58–0.98) amongst HIV-infected participants (**Figure 2**) than for detecting asymptomatic, subclinical TB (AUCs ranging from 0.59–0.86). However, because only two symptomatic HIV-infected and eight symptomatic HIV-uninfected participants with prevalent TB had available RNA, these exploratory results should be interpreted cautiously. As the majority of healthy participants and TB cases were asymptomatic, the AUC estimates for the pooled primary endpoint are very similar to estimates for the asymptomatic subgroup. Diagnostic performance for the secondary endpoint was not qualitatively different to that of the primary endpoint for all signatures in both HIV-uninfected and HIV-infected cohorts (**Figure S3-4**).

For the primary diagnostic endpoint and asymptomatic sub-group, no signatures met the minimal WHO TPP benchmarks for a triage test (sensitivity 90% and specificity 70%) in the HIV-uninfected cohort (**Figures 1 and Table S10**). However, five signatures (RISK11, RISK6, Roe1, Roe3 and Suliman4) met the optimal WHO TPP benchmark (sensitivity 95% and specificity 80%) amongst symptomatic participants, and the other four signatures (Francisco2, Maertzdorf4, Sweeney3, and Thompson5) either approached or met the minimal WHO TPP in this sub-group (**Figure 1**). Similarly amongst PLHIV, no signatures met the minimal WHO TPP benchmarks for a triage test (sensitivity 90% and specificity 70%) for the primary diagnostic endpoint and asymptomatic sub-group (**Figure 2 and Table S11**), however four signatures (RISK11, Maertzdorf4, Suliman4, and Thompson5) met the optimal WHO triage test benchmarks in participants with TB symptoms (**Figure 2**). Due to the small numbers of primary endpoint prevalent TB cases with evaluable transcriptomic signature scores amongst PLHIV (two symptomatic and six asymptomatic), precision around estimates was limited. For the primary diagnostic endpoint in the HIV-infected cohort, the upper limits of the 95% confidence intervals of performance metrics were within the bounds of the minimal WHO triage test criteria for most signatures, barring Francisco2, Sweeney3, and Thompson5 (**Table S11**).

Significant prognostic performance of signatures through 12 months in HIV-uninfected and 15-months in HIV-infected cohorts

We next evaluated transcriptomic signature prognostic performance for incident TB in all participants without primary endpoint prevalent TB at baseline who attended one or more follow-up visit. All signatures measured at enrolment had excellent performance for detecting HIV-uninfected participants who progressed to primary endpoint incident TB disease through six months follow-up, with AUCs ranging from 0.80–0.95 (**Figure 3 and Table S12**). Prognostic performance declined through 15 months follow-up (AUCs 0.49–0.66) with only RISK6 showing significant discrimination at this timepoint (AUC 0.66, 95% CI 0.51–0.82). Most signatures demonstrated good discrimination (AUCs 0.63–0.80) up to 12 months post enrolment. In the HIV-infected CORTIS-HR cohort, there were insufficient primary endpoint incident TB cases to evaluate prognostic performance prior to 15 months follow-up. Six of nine transcriptomic signatures were able to significantly differentiate between incident TB cases and controls in the HIV-infected cohort through 15 months, with AUCs ranging from 0.54–0.81 (**Figures 4 and Tables S13**). Lower 95% confidence bounds for the Francisco2, Sweeney3, and Thompson5 signatures crossed 0.50 indicating that prognostic performance at this timepoint was not significant. Prognostic performance through 15 months follow-up for the secondary endpoint was not qualitatively different to the primary endpoint in either cohort (**Figure S5-6**).

For primary endpoint incident TB, none of the signatures met the minimal WHO TPP benchmark for a prognostic test (sensitivity 75% and specificity 75%) in the HIV-uninfected cohort through 15 months follow-up (**Figure 3 and Table S12**); the Roe3 signature came closest at its optimal cut-point (sensitivity 70.0%, 95% CI 46.0–92.9; specificity 72.9%, 95% CI 71.1–74.7). However, through six months follow-up, six of eight parsimonious signatures met the minimal benchmark. The remaining two signatures, Suliman4 and Sweeney3, approached it with upper limits of the 95% confidence intervals for prognostic estimates exceeding the benchmark bounds (**Table S12**). Only RISK11 met the optimal WHO TPP benchmark for a prognostic test (sensitivity 90% and specificity 90%) through six months. Through 12 months follow-up, the upper limits of the 95% confidence intervals for all signatures exceeded the minimal WHO TPP benchmark, but only the Roe3 signature (sensitivity 84.9%, 95% CI 55.0–100; specificity 75.4%, 95% CI 73.7–77.1) met the benchmark. Amongst people living with HIV, only the Roe1 signature (sensitivity 83.8%, 95% CI 28.1–98.6; specificity 78.0%, 95% CI 74.5–81.0) met the minimal WHO TPP benchmark for a prognostic test through 15 months follow-up (**Figure 4 and Table S13**), however the upper 95% confidence interval bounds for six of eight of the other signatures (except Maertzdorf4 and Sweeney3) also met or approximated this benchmark.

Most TB signatures failed to differentiate participants with upper respiratory viral pathobionts from those with TB disease

Amongst participants with HIV we found that those without TB who had a detectable HIV plasma viral load (≥ 100 copies/mL) had significantly higher ($p < 0.0001$; or conversely lower for Roe1 and

Roe3) transcriptomic signature scores than those with suppressed HIV plasma viral load (<100 copies/mL). The exception was Thompson5, the only signature that does not include ISGs¹⁸, which was not affected by viral load ($p=0.87$; **Figure S7**). Among HIV-uninfected participants without TB, scores for most signatures (except Francisco2 and Sweeney3) differed significantly ($p<0.05$) between those with and without TB symptoms (**Figure 1**); we hypothesised that these differences were due to the presence of other comorbid infections in the symptomatic participants.

We tested this hypothesis in a cohort 1,000 HIV-uninfected CORTIS-01 participants who were consecutively enrolled into the respiratory pathobionts sub-study (**Figure S1C** and **Table S14**). All participants were investigated, irrespective of symptoms, for upper respiratory tract pathobionts in nasopharyngeal and oropharyngeal swabs using a multiplex bacteria and virus respiratory RT-qPCR panel. Viral and/or bacterial pathobionts were detected in 7.4% (74/1,000) and 38.9% (389/1,000) of participants respectively; with viral-bacterial co-detection in 3.3% (33/1,000) participants.

TB transcriptomic signature scores were significantly higher ($p<0.05$; or conversely lower for Roe1 and Roe3) in participants with any viral pathobiont, or both viral and bacterial pathobionts, compared to those with only bacterial, or no pathobionts, with the exception of Thompson5 (**Figure 5**). Most of the signatures significantly discriminated between viral and bacterial pathobionts, and viral and no pathobionts, but were unable to differentiate bacterial from no pathobionts. We also measured the Herberg2 signature²⁹, designed to discriminate between viral and bacterial infection in febrile children, in the respiratory pathobionts cohort; this signature performed equivalently to the other signatures in differentiating between the groups. Most of the upper respiratory bacteria detected (*H influenza* 21.3%, *S aureus* 10.2%, *S pneumonia* 9.2%, *M catarrhalis* 5.8%) commonly colonise the respiratory tract of healthy individuals, hence we are unable to determine the effect of pathogenic bacteria on signature scores. Thompson5 was the only signature for which no differences in scores were observed between participants with viral, bacterial, or no pathobionts, and subsequently was not able to differentiate between these groups (**Figure 5**).

28.6% (286/1,000) of participants were co-enrolled into the respiratory pathobionts sub-study and CORTIS-01 and consequently investigated for TB (**Figure S1C**). There were insufficient primary endpoint TB cases for meaningful analysis. However, 11 secondary endpoint (≥ 1 sputum sample positive for *Mtb*) prevalent TB cases (3.8%) were identified in CORTIS-01; 4/99 (4.0%) participants with bacteria detected in the upper respiratory tract and 7/150 (4.7%) with no pathobionts detected. No participants with viral upper respiratory pathobionts only ($n=41$), or mixed viral and bacterial pathobionts ($n=33$), had prevalent TB. The presence of bacterial upper respiratory pathobionts did not significantly affect TB diagnostic performance. Transcriptomic signatures differentiated between participants with prevalent TB and those with no pathobionts detected, with AUCs

500 between 0.63 and 0.79, and between participants with prevalent TB and those with other bacterial
 501 pathobionts with AUCs ranging from 0.60 to 0.81 (**Figure S8**). However, with the exception of
 502 Thompson5 (AUC 0.73, 95% CI 0.51–0.93), none of the other signatures were able to differentiate
 503 between participants with prevalent TB and those with viral upper respiratory pathobionts.

Discussion

We validated eight parsimonious TB transcriptomic signatures by microfluidic RT-qPCR and prospectively evaluated their diagnostic and prognostic accuracy in two large community screening cohorts of unselected HIV-uninfected and HIV-infected adults. We found that none of the signatures met the WHO triage test TPP benchmarks¹⁹ for subclinical TB, although performance for several signatures approached these thresholds. Parsimonious signature diagnostic performance in adults with TB symptoms was superior to that in asymptomatic participants, with most meeting or approaching the TPP benchmarks. In the HIV-uninfected cohort prognostic accuracy for incident TB within six months of testing was excellent for virtually all parsimonious signatures, while performance for TB within 12 months was moderate. Only one signature, RISK6, had any statistically significant prognostic ability for incident TB within 15 months of testing.

This waning accuracy mirrors the finding of a short prognostic window after signature testing in the meta-analysis by Gupta et al¹⁴, and supports the hypothesis that the timeframe of TB progression from infection to clinical disease is highly heterogeneous between individuals, ranging from months to years.³⁶⁻³⁸ We postulate that the declining performance distal to signature measurement may be due to low inflammatory activity in early disease stages, described as incipient TB. Individuals with incipient TB may spontaneously control the infection and thus not progress to disease³⁷, resulting in reduced specificity. In this regard, we suggest that the 2-year predictive window for a prognostic test proposed in the WHO framework²⁰ be revisited. Transient viral infections may also increase false positivity.¹⁷ Low sensitivity may reflect the high burden of population *Mtb* exposure and rate of new infections in this endemic setting.

Performance of parsimonious signatures in PLHIV appeared to match or surpass performance in participants without HIV, consistent with the findings of Turner and colleagues¹⁵. Most signatures were able to differentiate between those who would progress to TB disease (progressors) through 15 months follow-up and non-progressors in the HIV-infected cohort, and the Roe1 signature met the minimal WHO TPP benchmark for a prognostic test²⁰. The greater prognostic window of the signatures measured at baseline in PLHIV may reflect prevalent TB cases missed at study enrolment (possibly due to paucibacillary sputum), but detected further along the disease spectrum during follow-up. At each study visit, participants underwent symptom screening with symptom-guided TB investigations which would have missed asymptomatic disease at earlier timepoints. However, all participants provided sputum samples for TB investigation at the end-of-study visits, which may also explain the bulk of disease diagnosed at the month 15 visit. Most of the incident TB cases in the HIV-infected cohort were detected after 12 months follow-up; this may be due to the appearance of symptoms late in the disease course in immunosuppressed individuals, and as a result, incident TB cases may have been missed by symptom screening at earlier visits. It is also conceivable that the better performance in PLHIV is driven by ART-naïve individuals who have

increased risk of TB disease, but also have elevated signature scores (or lower for Roe1 and Roe3) due to their higher HIV plasma viral loads. However, 74% (142/193) of ART-naïve participants started ART during the study (**Table S9**), which would lower their risk of progression, and increase “false” positivity. Unfortunately our study was underpowered to evaluate signature performance stratified by viral load or ART status, however we postulate that higher signature scores in those with non-suppressed HIV plasma viral loads (**Figure S7**) would result in significantly lower specificity, but equivalent or higher sensitivity.

We recently showed that elevated RISK11 signature scores in healthy individuals without TB return to normal levels within three months irrespective of the use of TPT in the majority of individuals.¹⁷ The transient elevation of RISK11 may be due to self-clearance of *Mtb* infection in certain individuals, but is likely due to other exogenous factors that temporarily affect transcriptomic signature gene expression. Scores and performance of most parsimonious signatures, with the exception of the Thompson5 signature, were affected by the presence of upper respiratory viral, but not bacterial, pathobionts. Most transcriptomic signatures are primarily comprised of ISGs, which are known to be induced by viral infections, including HIV. The Thompson5 signature, which does not contain any ISGs, was unaffected by respiratory pathobionts or HIV. This supports the hypothesis that exogenous factors, and in particular viral infections, have a large impact on TB transcriptomic signature scores and performance due to induction of these genes; design of transcriptomic signatures which do not predominantly consist of ISGs,^{39,40} or which use a multinomial modelling⁴¹ approach to exclude viral infection, may address this shortcoming.

Given that this study was not designed to compare diagnostic and prognostic accuracy between signatures, and the 95% confidence intervals of performance metrics for most signatures were overlapping, we are hesitant to champion any particular signature. Several parsimonious signatures performed at least as well as RISK11. Performance of most signatures is promising for triage of symptomatic adults seeking care and PLHIV attending ART clinic, to rule out active disease and guide further investigation, and for short-term (6-12 months) prognosis of TB disease risk for targeted preventive therapy. Recent studies of 3HP have shown limited duration of efficacy (9 months) for biomarker-guided preventive therapy in HIV-uninfected adults in South Africa¹¹, and no additional benefit of universal annual retreatment of PLHIV on ART compared to a single course⁴². Our results further suggest that a prognostic window greater than 12 months after testing is unlikely to be of value in high transmission settings with associated risk of re-infection. This implies that annual re-testing would likely be necessary if such signatures are to be included in TB control strategies.⁴³ Parsimonious signature performance reported here warrants translation to point-of-care devices for clinical implementation studies.

No signatures met the WHO triage test TPP accuracy benchmark for diagnosing asymptomatic prevalent TB, which is a setback for community active case-finding efforts in endemic settings, such as South Africa, where there is a high burden of undiagnosed sub-clinical disease.^{2-5,11,12} However, we would note that these benchmarks were not intended for TB screening in a healthy community setting, but rather “to test persons with any symptoms or risk factors suggestive of TB who are seeking care”¹⁹, and thus the bar may have been too high for a strategy that seeks to find subclinical TB cases. Currently, chest radiography is the only sufficiently accurate screening tool in widespread use for detecting asymptomatic sub-clinical disease.²⁻⁵ Portable molecular screening has also been proposed as a possible universal screening tool for active case-finding,⁷ but this remains unaffordable in most settings. Given the recently recognised burden of asymptomatic TB in endemic settings and lack of affordable and accessible active case-finding tools to detect early and subclinical disease in community settings, we would propose that a new consensus framework for the development of screening tests for asymptomatic TB is urgently needed. Such a framework should also provide clear definitions of contentious concepts of asymptomatic “incipient” and “sub-clinical” TB.^{44,45} We have previously demonstrated a sequential inflammatory TB disease process^{22,36,46}. The trend of increasing signature scores from individuals without TB disease, through early or incipient (individuals who later progress to incident disease), sub-clinical, and clinical TB disease observed in this study (**Figures 1-2**) succinctly illustrates the stages of disease progression. Further trials are needed to determine whether earlier stages of TB progression require a full course of curative treatment, or an abbreviated regimen such as 3HP.

The correlation between most signatures (with the exception of Thompson5), and fact that no individual signature was significantly superior for diagnosis or prognosis, indicates they likely measure the same or similar biological pathways. Our study would suggest that performance of transcriptomic host-response signatures, based on discovery in active TB cases or progressor cohorts, has reached an apex and that further signature discovery using similar conventional methodology and study designs is unlikely to deliver significantly better diagnostic tools. What could be done differently? First, we would argue the necessity for recruiting larger heterogeneous unselected prospective discovery cohorts, with TB cases representing a spectrum from sub-clinical through clinical disease at prevalence rates that reflect the real epidemic, rather than carefully-curated case-control studies with uniform and highly enriched TB cases and homogenous healthy controls or inappropriate other diseases. In this study, all signatures performed superbly in the CTBC case-control cohort, but performance dipped substantially in the real-world CORTIS studies. Second, we have shown that parsimonious signatures perform at least as well as non-reduced signatures such as RISK11; early translation of parsimonious signatures to near point-of-care technology, such as RT-qPCR, would allow more rapid advancement along the developmental pipeline. Third, few studies have incorporated clinical features (such as age, sex, body-mass index, C-reactive protein, haemoglobin, monocyte-lymphocyte ratio, and HIV status or viral load)

into predictive host-response models, which may prove low-hanging fruit for improving signature performance.^{47,48} Finally, selection of genes which remain stable on repeat measurements through longitudinal follow-up in healthy individuals^{17,46}, and in individuals with common comorbid diseases and infections, may enhance predictive utility. Modular (genes selected from distinct biological pathways) or multinomial (multiple outcomes or covariates) modelling approaches to TB signature discovery may help to mitigate the effect of exogenous factors on signature performance.^{39,41}

A major strength of this study was the prospective recruitment of large community cohorts of unselected adults from multiple sites reflecting distinct patterns of *Mtb* exposure and rates of TB disease in an endemic setting. Despite intensive community active case-finding with enrichment of the HIV-uninfected study population for individuals at greatest risk of TB disease, sputum sample collection from all enrolled participants at enrolment and end-of-study visits, and symptom-triggered investigation during follow-up, there were relatively few prevalent and incident TB cases, limiting power for analysis of diagnostic and prognostic performance. We also did not investigate participants for extra-pulmonary TB, potentially missing cases amongst PLHIV. Another limitation was the high Suliman4 failure rate, resulting in less precise performance estimates for this signature. We also acknowledge that primer sequences selected for RT-qPCR may not be equivalent to those used for microarray or in the Cepheid Xpert MTB Host Response prototype cartridge^{49,50}. The large proportion of asymptomatic prevalent TB cases is a feature unique to this cohort, and may prove valuable for future biomarker discovery. Recruitment was restricted to South Africa, limiting generalisability of the results to other contexts; there is a need for similar large prospective diagnostic validation cohorts in other TB endemic regions.

Several parsimonious TB transcriptomic signatures hold promise for triage of symptomatic adults seeking care, screening of ART clinic attendees and other high-risk groups for further investigation, and prediction of short-term risk of TB for initiation of targeted preventive therapy. Point-of-care RT-qPCR testing platforms currently under development and in field implementation studies,^{49,50} would bring these signatures from the laboratory bench to the clinic bedside.

Author contributions

SCM, AP-N, AF-G, MH, and TJS conceived the study. AP-N, SKM, MH, and TJS implemented clinical studies, raised funds and/or provided the resources. MT, GW, KN, and GC were responsible for all site-level activities, including recruitment, clinical management, and data collection. SKM, MM, HM, MF, KH, and ME, and provided operational, technical, or laboratory support and project management. SCM, ME, and ON processed samples and performed the experiments. SCM, SKM, AF-G, and HM verified the underlying data. SCM and AF-G analysed data. SCM, SKM, TJS, and MH interpreted results and wrote the first draft of the manuscript. Members of the CTBC Study Team and the CORTIS Study Team performed clinical studies and processed samples. All authors had full access to the data, confirm the integrity of the data and its presentation, agree with its interpretation as discussed in the manuscript, and reviewed, revised, and approved the manuscript before submission. The corresponding author had final responsibility for the decision to submit for publication.

Competing interests statement

AP-N, GW, GC, TJS, and MH report grants from the Bill & Melinda Gates Foundation, during the conduct of the study; AP-N and GW report grants from the South African Medical Research Council, during the conduct of the study; GW and TJS report grants from the South African National Research Foundation, during the conduct of the study. In addition, AP-N and TJS have patents of the RISK11 and RISK6 signatures pending; GW has a patent “TB diagnostic markers” (PCT/IB2013/054377) issued and a patent “Method for diagnosing TB” (PCT/IB2017/052142) pending. All other authors had nothing to disclose.

Data availability

Deidentified signature scores, clinical metadata, and TB endpoint data will be made available with publication in the supplemental material (**Tables S15-18**). The public PCR probe dataset and metadata has been deposited in Zivahub (<https://doi.org/10.25375/uct.14999895>), an open access data repository hosted by the University of Cape Town’s institutional data repository powered by Figshare for Institutions.

Funding

The CTBC study was funded by the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council (SAMRC) with funds received from the South African Department of Science and Technology. The CORTIS studies were funded by the Bill and Melinda Gates Foundation (BMGF; OPP1116632, OPP1137034, and OPP1151915) and also the SHIP Unit of the SAMRC with funds received from the South African Department of Science and Technology. Respiratory infection assays were funded by the NIH (AI123780).

683 SCM is a recipient of PhD funding from the South African Medical Association (SAMA), the Fogarty
684 International Center of the National Institutes of Health (NIH) under Award Number D43
685 TW010559, the Harry Crossley Clinical Research Fellowship, and the South African Medical
686 Research Council (SAMRC) through its Division of Research Capacity Development under the
687 SAMRC Clinician Researcher Programme. The content is solely the responsibility of the authors
688 and does not necessarily represent the official views of SAMA, the NIH, the Harry Crossley
689 Foundation, or the SAMRC. The funders had no role in study design, data collection and analysis,
690 decision to publish, or preparation of the manuscript.

References

1. WHO. Global tuberculosis report 2020. Geneva: World Health Organization, 2020. https://www.who.int/tb/publications/global_report/en/ (accessed October 15, 2020).
2. Onozaki I, Law I, Sismanidis C, Zignol M, Glaziou P, Floyd K. National tuberculosis prevalence surveys in Asia, 1990-2012: an overview of results and lessons learned. *Tropical Medicine and International Health*. 2015;20(9):1128-45. doi: 10.1111/tmi.12534.
3. Frascella B, Richards AS, Sossen B, Emery JC, Odone A, Law I, et al. Subclinical tuberculosis disease - a review and analysis of prevalence surveys to inform definitions, burden, associations and screening methodology. *Clinical Infectious Diseases*. 2020. doi: 10.1093/cid/ciaa1402.
4. Law I, Floyd K, African TBPSG. National tuberculosis prevalence surveys in Africa, 2008-2016: an overview of results and lessons learned. *Tropical Medicine and International Health*. 2020;25(11):1308-27. doi: 10.1111/tmi.13485.
5. South African National Department of Health. The First National TB Prevalence Survey: South Africa 2018. 2021. www.knowledgehub.org.za/elibrary/first-national-tb-prevalence-survey-south-africa-2018 (accessed May 22, 2021).
6. Harris M, Qi A, Jeagal L, Torabi N, Menzies D, Korobitsyn A, et al. A systematic review of the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest x-rays for pulmonary tuberculosis. *PLoS ONE*. 2019;14(9):e0221339. doi: 10.1371/journal.pone.0221339.
7. Calligaro GL, Zijenah LS, Peter JG, Theron G, Buser V, McNerney R, et al. Effect of new tuberculosis diagnostic technologies on community-based intensified case finding: a multicentre randomised controlled trial. *Lancet Infectious Diseases*. 2017;17(4):441-50. doi: 10.1016/S1473-3099(16)30384-X.
8. Fiore-Gartland A, Carpp LN, Naidoo K, Thompson E, Zak DE, Self S, et al. Considerations for biomarker-targeted intervention strategies for tuberculosis disease prevention. *Tuberculosis (Edinb)*. 2018;109:61-8. doi: 10.1016/j.tube.2017.11.009.
9. Darboe F, Mbandi SK, Thompson EG, Fisher M, Rodo M, van Rooyen M, et al. Diagnostic performance of an optimized transcriptomic signature of risk of tuberculosis in cryopreserved peripheral blood mononuclear cells. *Tuberculosis (Edinb)*. 2018;108:124-6. doi: 10.1016/j.tube.2017.11.001.
10. Darboe F, Mbandi SK, Naidoo K, Yende-Zuma N, Lewis L, Thompson EG, et al. Detection of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. *Front Microbiol*. 2019;10:1441. doi: 10.3389/fmicb.2019.01441.

11. Scriba TJ, Fiore-Gartland A, Penn-Nicholson A, Mulenga H, Kimbung Mbandi S, Borate B, et al. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled trial. *Lancet Infectious Diseases*. 2021;21(3):354-65. doi: 10.1016/S1473-3099(20)30914-2.
12. Mendelsohn SC, Fiore-Gartland A, Penn-Nicholson A, Mulenga H, Mbandi SK, Borate B, 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(6):e841-e53. doi: 10.1016/S2214-109X(21)00045-0.
13. Warsinske H, Vashisht R, Khatri P. Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. *PLoS Medicine*. 2019;16(4):e1002786. doi: 10.1371/journal.pmed.1002786.
14. Gupta RK, Turner CT, Venturini C, Esmail H, Rangaka MX, Copas A, et al. Concise whole blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-level pooled meta-analysis. *Lancet Respir Med*. 2020;8(4):395-406. doi: 10.1016/S2213-2600(19)30282-6.
15. Turner CT, Gupta RK, Tsaliki E, Roe JK, Mondal P, Nyawo GR, et al. Blood transcriptional biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, observational, diagnostic accuracy study. *Lancet Respir Med*. 2020;8(4):407-19. doi: 10.1016/S2213-2600(19)30469-2.
16. Mulenga H, Zaichenberger CZ, Bunyasi EW, Mbandi SK, Mendelsohn SC, Kagina B, et al. Performance of diagnostic and predictive host blood transcriptomic signatures for Tuberculosis disease: A systematic review and meta-analysis. *PLoS ONE*. 2020;15(8):e0237574. doi: 10.1371/journal.pone.0237574.
17. Mulenga H, Musvosvi M, Mendelsohn SC, Penn-Nicholson A, Mbandi SK, Fiore-Gartland A, et al. Longitudinal dynamics of a blood transcriptomic signature of tuberculosis. 2021 (Submitted).
18. Thompson EG, Du Y, Malherbe ST, Shankar S, Braun J, Valvo J, et al. Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis (Edinb)*. 2017;107:48-58. doi: 10.1016/j.tube.2017.08.004.
19. WHO. High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. Geneva: World Health Organization, 2014. www.who.int/tb/publications/tpp_report/en/ (accessed May 22, 2020).
20. WHO. 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. Geneva: World Health Organization, 2017. <http://apps.who.int/iris/handle/10665/259176> (accessed May 22, 2020).
21. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527. doi: 10.1136/bmj.h5527.

22. Penn-Nicholson A, Mbandi SK, Thompson E, Mendelsohn SC, Suliman S, Chegou NN, et al. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response. *Sci Rep.* 2020;10(1):8629. doi: 10.1038/s41598-020-65043-8.
23. Francisco NM, Fang YM, Ding L, Feng S, Yang Y, Wu M, et al. Diagnostic accuracy of a selected signature gene set that discriminates active pulmonary tuberculosis and other pulmonary diseases. *Journal of Infection.* 2017;75(6):499-510. doi: 10.1016/j.jinf.2017.09.012.
24. Maertzdorf J, McEwen G, Weiner J, 3rd, Tian S, Lader E, Schriek U, et al. Concise gene signature for point-of-care classification of tuberculosis. *EMBO Molecular Medicine.* 2016;8(2):86-95. doi: 10.15252/emmm.201505790.
25. Suliman S, Thompson EG, Sutherland J, Weiner J, 3rd, Ota MOC, Shankar S, et al. Four-Gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *American Journal of Respiratory and Critical Care Medicine.* 2018;197(9):1198-208. doi: 10.1164/rccm.201711-2340OC.
26. Roe JK, Thomas N, Gil E, Best K, Tsaliki E, Morris-Jones S, et al. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. *JCI Insight.* 2016;1(16):e87238. doi: 10.1172/jci.insight.87238.
27. Roe J, Venturini C, Gupta RK, Gurry C, Chain BM, Sun Y, et al. Blood Transcriptomic Stratification of Short-term Risk in Contacts of Tuberculosis. *Clinical Infectious Diseases.* 2020;70(5):731-7. doi: 10.1093/cid/ciz252.
28. Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med.* 2016;4(3):213-24. doi: 10.1016/S2213-2600(16)00048-5.
29. Herberg JA, Kaforou M, Wright VJ, Shailes H, Eleftherohorinou H, Hoggart CJ, et al. Diagnostic Test Accuracy of a 2-Transcript Host RNA Signature for Discriminating Bacterial vs Viral Infection in Febrile Children. *JAMA.* 2016;316(8):835-45. doi: 10.1001/jama.2016.11236.
30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC Bioinformatics.* 2011;12:77. doi: 10.1186/1471-2105-12-77.
31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics.* 1988;44(3):837-45. doi: 10.2307/2531595.
32. Zheng Y, Cai T, Brown M. Package 'survAccuracyMeasures'. 2014. <http://mdbrown.github.io/survAccuracyMeasures/> (accessed May 22, 2020).
33. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two criteria based on the receiver operating characteristic curve. *American Journal of Epidemiology.* 2006;163(7):670-5. doi: 10.1093/aje/kwj063.

34. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide for medical statisticians. *Statistics in Medicine*. 2000;19(9):1141-64. doi: 10.1002/(sici)1097-0258(20000515)19:9<1141::Aid-sim479>3.0.Co;2-f.
35. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal statistical society: series B (Methodological)*. 1995;57(1):289-300.
36. Scriba TJ, Penn-Nicholson A, Shankar S, Hraha T, Thompson EG, Sterling D, et al. Sequential inflammatory processes define human progression from *M. tuberculosis* infection to tuberculosis disease. *PLoS Pathogens*. 2017;13(11):e1006687. doi: 10.1371/journal.ppat.1006687.
37. Drain PK, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, et al. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. *Clinical Microbiology Reviews*. 2018;31(4):e00021-18. doi: 10.1128/CMR.00021-18.
38. Scriba TJ, Mendelsohn SC. Headway made towards biosignatures for incipient tuberculosis. *Lancet Respir Med*. 2020;8(4):328-30. doi: 10.1016/S2213-2600(19)30355-8.
39. Singhanian A, Verma R, Graham CM, Lee J, Tran T, Richardson M, et al. A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nature Communications*. 2018;9(1):2308. doi: 10.1038/s41467-018-04579-w.
40. Esmail H, Lai RP, Lesosky M, Wilkinson KA, Graham CM, Horswell S, et al. Complement pathway gene activation and rising circulating immune complexes characterize early disease in HIV-associated tuberculosis. *Proceedings of the National Academy of Sciences of the United States of America*. 2018;115(5):E964-E73. doi: 10.1073/pnas.1711853115.
41. Duffy FJ, Thompson EG, Scriba TJ, Zak DE. Multinomial modelling of TB/HIV co-infection yields a robust predictive signature and generates hypotheses about the HIV+TB+ disease state. *PLoS ONE*. 2019;14(7):e0219322. doi: 10.1371/journal.pone.0219322.
42. Churchyard G, Cardenas V, Chihota V, Mngadi K, Sebe M, Brumskine W, et al. Effectiveness of 3HP annually vs once for HIV-positive people: The WHIP3TB Trial [CROI Abstract 143]. In Special Issue: Abstracts From the 2020 Conference on Retroviruses and Opportunistic Infections. *Top Antivir Med*. 2020;28(1).
43. Sumner T, Mendelsohn SC, Scriba TJ, Hatherill M, White R. The impact of blood transcriptomic biomarker targeted tuberculosis preventive therapy in people living with HIV: a mathematical modelling study. 2021 (Submitted).
44. Kendall EA, Shrestha S, Dowdy DW. The Epidemiological Importance of Subclinical Tuberculosis. A Critical Reappraisal. *American Journal of Respiratory and Critical Care Medicine*. 2021;203(2):168-74. doi: 10.1164/rccm.202006-2394PP.
45. Pierce CW. Subclinical Tuberculosis: Some Flies in the Ointment. *American Journal of Respiratory and Critical Care Medicine*. 2021. doi: 10.1164/rccm.202101-0227LE.

46. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *Lancet*. 2016;387(10035):2312-22. doi: 10.1016/S0140-6736(15)01316-1.
47. Sivakumaran D, Jenum S, Vaz M, Selvam S, Ottenhoff THM, Haks MC, et al. Combining host-derived biomarkers with patient characteristics improves signature performance in predicting tuberculosis treatment outcomes. *Commun Biol*. 2020;3(1):359. doi: 10.1038/s42003-020-1087-x.
48. Mendelsohn SC, Mbandi SK, Hatherill M, Scriba TJ. Blood transcriptional signatures for tuberculosis testing. *Lancet Respir Med*. 2020;8(4):330-1. doi: 10.1016/S2213-2600(20)30045-X.
49. Sodersten E, Ongarello S, Mantsoki A, Wyss R, Persing DH, Banderby S, et al. Diagnostic Accuracy Study of a Novel Blood-Based Assay for Identification of Tuberculosis in People Living with HIV. *Journal of Clinical Microbiology*. 2021;59(3). doi: 10.1128/JCM.01643-20.
50. Moreira FMF, Verma R, Pereira Dos Santos PC, Leite A, da Silva Santos A, de Araujo RCP, et al. Blood-based host biomarker diagnostics in active case finding for pulmonary tuberculosis: A diagnostic case-control study. *EClinicalMedicine*. 2021;33:100776. doi: 10.1016/j.eclinm.2021.100776.

Figure legends

Figure 1: Parsimonious signature diagnostic performance for prevalent TB and signature score distributions in people without HIV

Paired receiver operating characteristic (ROC) curves and box-and-whisker plots for the (A-B) Darboe11 (RISK11), (C-D) Francisco2, (E-F) Maertzdorf4 (DIAG4), (G-H) Penn-Nicholson6 (RISK6), (I-J) Roe1 (BATF2), (K-L) Roe3, (M-N) Suliman4 (RISK4), (O-P) Sweeney3, and (Q-R) Thompson5 (RESPONSE5) signatures in the CORTIS-01 study of people without HIV. The ROC curves depict diagnostic performance (area under the curve, AUC, with 95% CI) of the parsimonious signatures for the primary endpoint (1' Endpoint), i.e. TB diagnosed on two or more liquid culture-positive or Xpert MTB/RIF-positive sputum samples. The ROC curves show participants with symptomatic clinical prevalent TB versus symptomatic controls (Symptom+), and participants with asymptomatic, subclinical prevalent TB versus asymptomatic controls (Symptom-). The shaded areas represent the 95% CIs. The solid box depicts the optimal criteria (95% sensitivity and 80% specificity) and the dashed box depicts the minimal criteria (90% sensitivity and 70% specificity) set out in the WHO Target Product Profile for a triage test.¹⁹ The box-and-whisker plots depict signature score (measured at enrolment) distribution by symptom status (each dot represents a participant) in asymptomatic and symptomatic participants with no TB, participants who progressed to incident TB, and participants with prevalent subclinical (asymptomatic) and clinical (symptomatic) TB. Prevalent and incident TB comprised all primary endpoint cases. Symptoms were recorded at the time of enrolment for participants without TB and those with prevalent TB. p values for comparison of median signature scores between groups in box-and-whisker plots were calculated with the Mann-Whitney *U* test and corrected for multiple comparisons by use of the Benjamini-Hochberg Procedure.³⁵ Boxes depict the IQR, the midline represents the median, and the whiskers indicate the IQR $\pm$ (1.5 $\times$ IQR). (S) Summary of signature diagnostic performance in order of primary endpoint AUC estimates. The diagnostic AUC estimates in symptomatic and asymptomatic participant sub-groups are also shown. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, and the black dotted line indicates the lower bound of the 95% CI for the best performing signature for the primary endpoint.

Figure 2: Parsimonious signatures diagnostic performance for prevalent TB and signature score distributions in people living with HIV.

Paired receiver operating characteristic (ROC) curves and box-and-whisker plots for the (A-B) Darboe11 (RISK11), (C-D) Francisco2, (E-F) Maertzdorf4 (DIAG4), (G-H) Penn-Nicholson6 (RISK6), (I-J) Roe1 (BATF2), (K-L) Roe3, (M-N) Suliman4 (RISK4), (O-P) Sweeney3, and (Q-R) Thompson5 (RESPONSE5) signatures in the CORTIS-HR study of people living with HIV.

The ROC curves depict diagnostic performance (area under the curve, AUC, with 95% CI) of the parsimonious signatures for TB diagnosed for the primary endpoint (1' Endpoint), i.e. TB diagnosed on two or more liquid culture-positive or Xpert MTB/RIF-positive sputum samples. The ROC curves show participants with symptomatic clinical prevalent TB versus symptomatic controls (Symptom+), and participants with asymptomatic subclinical prevalent TB versus asymptomatic controls (Symptom-). The shaded areas represent the 95% CIs. The solid box depicts the optimal criteria (95% sensitivity and 80% specificity) and the dashed box depicts the minimal criteria (90% sensitivity and 70% specificity) set out in the WHO Target Product Profile for a triage test.¹⁹

The box-and-whisker plots depict signature score (measured at enrolment) distribution by symptom status (each dot represents a participant) in asymptomatic and symptomatic participants with no TB, participants who progressed to incident TB, and participants with prevalent subclinical (asymptomatic) and clinical (symptomatic) TB. Prevalent and incident TB comprised all primary endpoint cases. Symptoms were recorded at the time of enrolment for participants without TB and those with prevalent TB. p values for comparison of median signature scores between groups in box-and-whisker plots were calculated with the Mann-Whitney *U* test and corrected for multiple comparisons by use of the Benjamini-Hochberg Procedure.³⁵ Boxes depict the IQR, the midline represents the median, and the whiskers indicate the IQR $\pm$ (1.5 $\times$ IQR).

(S) Summary of signature diagnostic performance in order of primary endpoint AUC estimates. The diagnostic AUC estimates in symptomatic and asymptomatic participant sub-groups are also shown. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, and the black dotted line indicates the lower bound of the 95% CI for the best performing signature for the primary endpoint. The black dashed line (Figure 2S only) indicates an AUC cut-off of 0.5.

Figures 3: Parsimonious signature prognostic performance for incident TB in people without HIV

Receiver operating characteristic curves depicting prognostic performance (area under the curve, AUC, with 95% CI) of the (A) Darboe11 (RISK11), (B) Francisco2, (C) Maertzdorf4 (DIAG4), (D) Penn-Nicholson6 (RISK6), (E) Roe1 (BATF2), (F) Roe3, (G) Suliman4 (RISK4), (H) Sweeney3, and (I) Thompson5 (RESPONSE5) signatures for incident TB diagnosed on two or more liquid culture-positive or Xpert MTB/RIF-positive sputum samples (primary endpoint) through 6, 12, and 15 months follow-up in the CORTIS-01 study of people without HIV.

The shaded areas represent the 95% CIs. The solid box depicts the optimal criteria (90% sensitivity and 90% specificity) and the dashed box depicts the minimal criteria (75% sensitivity and 75% specificity) set out in the WHO Target Product Profile for an incipient TB test.²⁰

(J) Summary of signature prognostic performance in order of primary endpoint AUC estimates through 12 months follow-up. The prognostic AUC estimates through 6 and 15 months are also shown. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, and the black dashed line indicates an AUC cut-off of 0.5.

Figures 4: Parsimonious signature prognostic performance for prevalent TB in people living with HIV

Receiver operating characteristic curves depicting prognostic performance (area under the curve, AUC, with 95% CI) of the (A) Darboe11 (RISK11), (B) Francisco2, (C) Maertzdorf4 (DIAG4), (D) Penn-Nicholson6 (RISK6), (E) Roe1 (BATF2), (F) Roe3, (G) Suliman4 (RISK4), (H) Sweeney3, and (I) Thompson5 (RESPONSE5) signatures for incident TB diagnosed on two or more liquid culture-positive or Xpert MTB/RIF-positive sputum samples (primary endpoint) through 15 months follow-up in the CORTIS-HR study of people living with HIV.

The shaded areas represent the 95% CIs. The solid box depicts the optimal criteria (90% sensitivity and 90% specificity) and the dashed box depicts the minimal criteria (75% sensitivity and 75% specificity) set out in the WHO Target Product Profile for an incipient TB test.²⁰ (J) Summary of signature prognostic performance in order of primary endpoint AUC estimates through 15 months follow-up. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, the red dotted line indicates the lower bound of the 95% CI for the best performing signature for the primary endpoint, and the black dashed line indicates an AUC cut-off of 0.5.

Figures 5: Parsimonious signature performance for differentiating participants with viral or bacterial upper respiratory tract pathobionts, and those without any pathobionts

Paired receiver operating characteristic (ROC) curves and box-and-whisker plots for the (A-B) Darboe11 (RISK11), (C-D) Francisco2, (E-F) Herberg2, (G-H) Maertzdorf4 (DIAG4), (I-J) Penn-Nicholson6 (RISK6), (K-L) Roe1 (BATF2), (M-N) Roe3, (O-P) Suliman4 (RISK4), (Q-R) Sweeney3, and (S-T) Thompson5 (RESPONSE5) signatures in all participants without prevalent TB in the respiratory pathobionts sub-study (11 participants with secondary endpoint prevalent TB excluded).

The box-and-whisker plots depict signature score distributions in participants with no upper respiratory pathobionts (n=563), bacterial upper respiratory pathobionts only (n=352), viral upper respiratory pathobionts only (n=41), both viral and bacterial upper respiratory pathobionts (n=33). p values for comparison of median signature scores between groups in box-and-whisker plots were calculated with the Mann-Whitney *U* test and corrected for multiple comparisons by use of the Benjamini-Hochberg Procedure.³⁵ Only p-values below 0.1 are shown. Each dot represents a participant. Boxes depict the IQR, the midline represents the median, and the whiskers indicate the $IQR \pm (1.5 \times IQR)$.

The ROC curves depict performance (area under the curve, AUC, with 95% CI) of the parsimonious signatures in differentiating between participants with viral upper respiratory pathobionts and participants with bacterial upper respiratory pathobionts, between participants with viral pathobionts and those with no pathobionts, or between participants with bacterial pathobionts and those with no pathobionts. The shaded areas represent the 95% CIs.

Figure 1

Figure 2

Figure 3

Figure 4

Figure 5