



OPEN

RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment response

Adam Penn-Nicholson^{1,3,4}, Stanley Kimbung Mbandi^{1,3,4}, Ethan Thompson^{2,3,4}, Simon C. Mendelsohn^{1,3,4}, Sara Suliman^{1,3}, Novel N. Chegou⁴, Stephanus T. Malherbe⁴, Fatoumatta Darboe¹, Mzwandile Erasmus¹, Willem A. Hanekom¹, Nicole Bilek¹, Michelle Fisher¹, Stefan H. E. Kaufmann^{5,6}, Jill Winter⁷, Melissa Murphy¹, Robin Wood⁸, Carl Morrow⁸, Ildiko Van Rhijn³, Branch Moody³, Megan Murray⁹, Bruno B. Andrade¹⁰, Timothy R. Sterling¹¹, Jayne Sutherland¹², Kogieleum Naidoo^{13,14}, Nesri Padayatchi^{13,14}, Gerhard Walzl⁴, Mark Hatherill¹, Daniel Zak², Thomas J. Scriba¹✉, The Adolescent Cohort Study team*, The GC6-74 Consortium*, The SATVI Clinical and Laboratory Team*, The ScreenTB Consortium*, The AE-TBC Consortium*, The RePORT Brazil Team*, Peruvian Household Contacts Cohort Team* & The CAPRISA IMPRESS team*

Improved tuberculosis diagnostics and tools for monitoring treatment response are urgently needed. We developed a robust and simple, PCR-based host-blood transcriptomic signature, RISK6, for multiple applications: identifying individuals at risk of incident disease, as a screening test for subclinical or clinical tuberculosis, and for monitoring tuberculosis treatment. RISK6 utility was validated by blind prediction using quantitative real-time (qRT) PCR in seven independent cohorts. Prognostic performance significantly exceeded that of previous signatures discovered in the same cohort. Performance for diagnosing subclinical and clinical disease in HIV-uninfected and HIV-infected persons, assessed by area under the receiver-operating characteristic curve, exceeded 85%. As a screening test for tuberculosis, the sensitivity at 90% specificity met or approached the benchmarks set out in World Health Organization target product profiles for non-sputum-based tests. RISK6 scores correlated with lung immunopathology activity, measured by positron emission tomography, and tracked treatment response, demonstrating utility as treatment response biomarker, while predicting treatment failure prior to treatment initiation. Performance of the test in capillary blood samples collected by finger-prick was noninferior to venous blood collected in PAXgene tubes. These results support incorporation of RISK6 into rapid, capillary blood-based point-of-care PCR devices for prospective assessment in field studies.

¹South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular Medicine and Division of Immunology, Department of Pathology, University of Cape Town, Cape Town, South Africa. ²Center for Infectious Disease Research, Seattle, WA, USA. ³Brigham and Women's Hospital, Division of Rheumatology, Immunity and Inflammation, Harvard Medical School, Boston, USA. ⁴DST-NRF Centre of Excellence for Biomedical Tuberculosis Research; South African Medical Research Council Centre for Tuberculosis Research; Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Stellenbosch University, Cape Town, South Africa. ⁵Max Planck Institute for Infection Biology, Berlin, Germany. ⁶Hagler Institute for Advanced Study at Texas A&M University, College Station, TX, USA. ⁷Catalysis Foundation for Health, San Ramon, CA, USA. ⁸Desmond Tutu HIV Centre, and Institute of Infectious Disease and Molecular Medicine (IDM), University of Cape Town, Cape Town, South Africa. ⁹Department of Global Health and Social Medicine, and Division of Global Health Equity, Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA. ¹⁰Instituto Gonçalo Moniz, Fundação Oswaldo Cruz, Salvador, Brazil. ¹¹Division of Infectious Diseases, Department of Medicine, Vanderbilt University School of Medicine, Nashville, USA. ¹²Vaccines and Immunity, Medical Research Council Unit, Fajara, The Gambia. ¹³Centre for the AIDS Programme of Research in Africa, Durban, South Africa. ¹⁴South African Medical Research Council-CAPRISA HIV-TB Pathogenesis and Treatment Research Unit, Durban, South Africa. ^{3,4}These authors contributed equally: Adam Penn-Nicholson, Stanley Kimbung Mbandi, Ethan Thompson and Simon C. Mendelsohn. *Lists of authors and their affiliations appear at the end of the paper. ✉e-mail: thomas.scriba@uct.ac.za

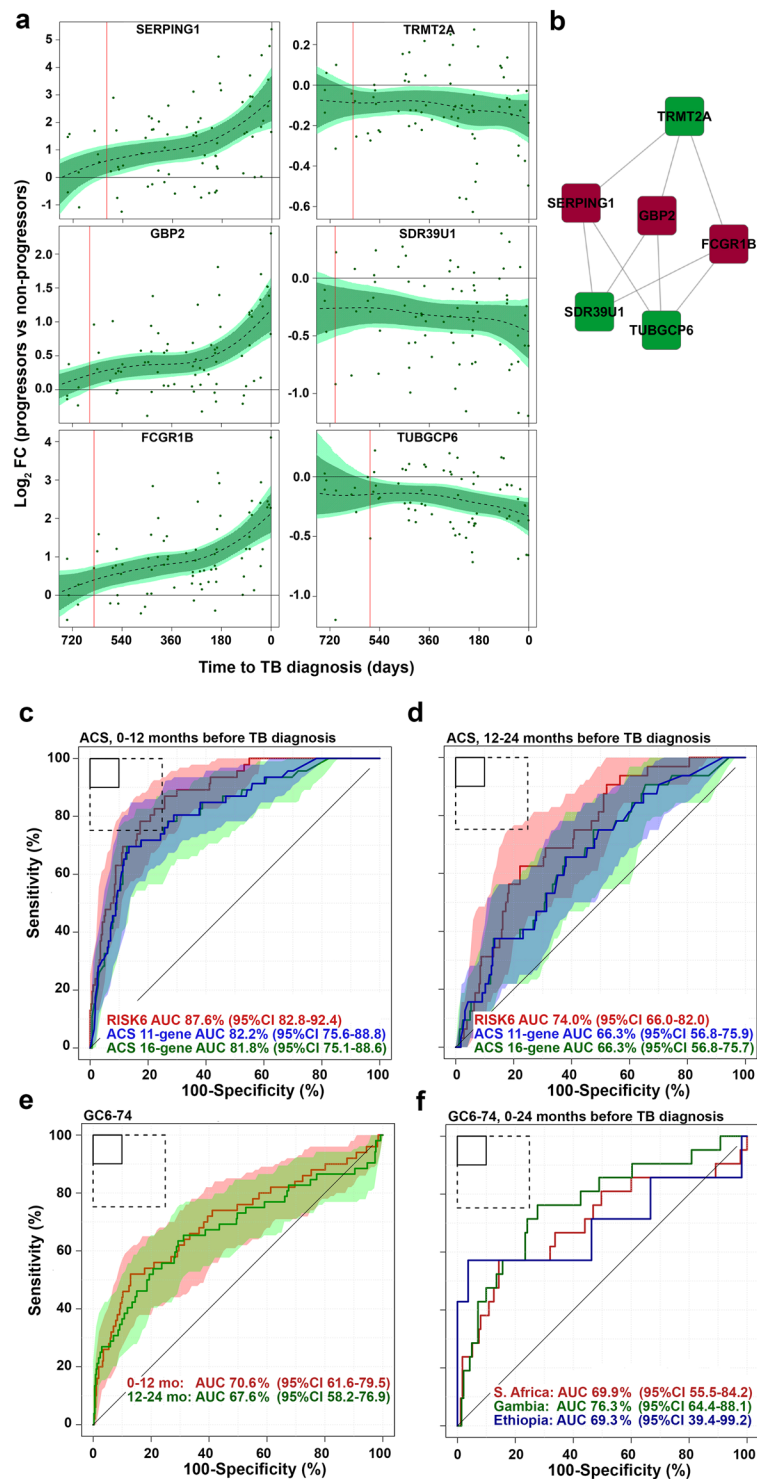


Figure 1. Discovery of the RISK6 signature. **(a)** Expression kinetics of the six transcripts in RISK6 signature over time, measured by RNA-sequencing and expressed as log₂ fold change between matched adolescent progressors and controls and modelled as non-linear splines (dotted lines). Light green shading represents 95% CI for the temporal trends, computed by performing 2000 spline fitting iterations after bootstrap resampling from the full dataset. Transcripts that are upregulated during TB progression are on the left and those that are downregulated on the right. **(b)** RISK6 comprises nine pairs that each link a transcript that is upregulated during TB progression with one that is downregulated, relative to healthy controls. Lines indicate pairing in progressors are in red nodes and those that are downregulated in progressors are in green. **(c and d)** Receiver operating characteristic (ROC) curves depicting the performance (model fit) of RISK6, relative to the ACS 11-gene⁷ and ACS 16-gene⁶ signatures, measured by qRT-PCR on RNA from whole blood samples collected from participants of the ACS cohort within one year of TB disease diagnosis **(c)**, or 1–2 years before TB

diagnosis (d). Shaded areas depict the 95% CI. (e) Receiver operating characteristic (ROC) curves depicting prognostic performance, by blind prediction of RISK6, measured by qRT-PCR on RNA from whole blood RNA collected from participants of the GC6-74 cohort of household contacts. Shaded areas depict the 95% CI. (f) Prognostic performance of RISK6 for incident TB in household contacts from South Africa, The Gambia or Ethiopia, measured by qRT-PCR on RNA from whole blood RNA collected within 1 year of TB diagnosis from participants of the GC6-74 cohort. The boxes in the top left corner represent the optimal (solid line) and minimum (dotted line) criteria set out in the target product profile for an incipient TB test. Performance on 0–2 years before TB diagnosis is shown in Table 1.

The “End Tuberculosis Strategy” of the World Health Organization (WHO) aims to reduce the annual incidence of tuberculosis (TB) to less than 10 cases per 100,000 people by the year 2035¹. To achieve this goal the primary proposed strategy is to increase and improve efforts to find and treat individuals with active TB disease, to conduct universal screening of those at high risk, and to provide preventive therapy to those at risk of progressing to active TB disease¹. There is thus a need for improved prognostic and diagnostic tests to identify those at risk of incident TB and those with subclinical or active TB, for appropriate treatment. The provision and management of TB treatment, as well as monitoring a patient’s response to treatment, also require much improvement. The standard 6-month regimen of treatment appears to be unnecessarily long for many patients with drug-susceptible TB, while insufficient to cure some patients, even in clinical trials when treatment adherence is maximized². Experimental regimens tested in recent clinical trials have also been inadequate to cure treatment-refractory patients³. Collectively, these data support the now accepted principle that TB exists in a pathophysiological spectrum that spans several stages of infection, subclinical and active disease, including distinct stages of treatment outcome. Achieving the “End Tuberculosis Strategy” clearly depends on approaches that can place an individual into the stage of this spectrum such that clinical management is appropriate.

A universal, non-sputum biomarker capable of predicting progression to active TB, diagnosing disease and monitoring the response to TB treatment would be a major advance in the efforts to achieve the “End Tuberculosis Strategy”. We hypothesized that a single, parsimonious host-blood transcriptomic signature can be developed for all three purposes with performance criteria that meet the target product profiles for tests to predict TB progression⁴ and for a TB screening test⁵ proposed by the WHO.

We sought to discover and validate a parsimonious and robust blood transcriptomic signature with applicability for predicting incident TB, as a triage test for identifying those who should be further investigated for TB disease, and for monitoring of TB treatment response. We explicitly set out to develop this signature for ultimate translation to a hand-held point-of-care platform and therefore performed all analyses, including signature training and performance assessments in all validation cohorts, by quantitative RT-PCR, using a highly standardized protocol and locked-down analysis algorithm. We assessed performance of RISK6 by blind prediction as a prognostic test for incident TB, as a TB diagnostic in HIV-uninfected and HIV-infected individuals, including individuals presenting with symptoms requiring investigation for TB at primary health care centres, and as a treatment response biomarker. We also tested the robustness of RISK6 and report performance of RISK6 measured in capillary blood samples collected by finger-prick, facilitating the way for incorporation into point-of-care diagnostic devices.

Results

Prognostic performance of RISK6 in the adolescent cohort study discovery cohort. The RISK6 signature was discovered on samples from adolescent progressors and controls (Supplementary Figure 1a) by selecting the smallest set of transcripts with the best prognostic performance based on qRT-PCR data. RISK6 comprises an ensemble of 9 transcript pairs formed between three transcripts upregulated in progressors (GBP2, FCGR1B, and SERPING1), and three transcripts downregulated in progressors (TUBGCP6, TRMT2A, and SDR39U1), relative to non-progressors (Supplementary Table 1 and Fig. 1a,b). We first sought to determine if the prognostic performance of RISK6 for incident TB in the discovery cohort was comparable to that of the previously published ACS 16-gene signature⁶, consisting of 57 transcripts (PCR primer/probe assays) or the ACS 11-gene version (48 PCR primer/probe assays), which was developed for greater throughput in multiplex assays⁷. The PCR-based RISK6 and both ACS signatures readily discriminated between Adolescent Cohort Study progressor and non-progressor samples collected within 12 months of TB diagnosis (Fig. 1c). Interestingly, prognostic performance of RISK6, estimated by model fit (AUC 87.6%, 95%CI 82.8–92.4), was significantly better than ACS 16-gene (AUC 81.8%, 95%CI 75.1–88.6, DeLong test $p = 0.024$) and ACS 11-gene (AUC 82.2%, 95%CI 75.6–88.8, DeLong test $p = 0.03$). As observed previously with the 16-gene signature⁶, RISK6 also discriminated between progressors and non-progressors using samples collected between 12 and 24 months before TB disease diagnosis (AUC 74.0%, 95%CI 66.0–82.0), although discrimination was weaker than observed on samples within a year of diagnosis (Fig. 1d).

Validation of RISK6 prognostic performance in the GC6–74 cohort. We validated prognostic performance of the RISK6 signature for incident TB by blind prediction on the independent GC6–74 cohorts of household TB contacts from South Africa, The Gambia and Ethiopia, who either progressed to TB or remained asymptomatic⁸. RISK6 significantly discriminated between GC6-74 progressors and non-progressors on samples collected within 12 months of incident TB diagnosis (AUC 70.6%, 95%CI 61.6–79.5) and those collected 12–24 months before TB diagnosis (AUC 67.6%, 95%CI 58.2–76.9, Fig. 1e and Table 1). At a sensitivity threshold of 75%, RISK6 achieved a specificity of 50.3% within 1 year of diagnosis in the GC6-74 cohort, which does not meet the WHO target product profile (TPP) for a test predicting progression from tuberculosis infection to active disease⁴ (Table 2).

	Gambia		South Africa		Ethiopia	
	AUC (95% CI)	p	AUC (95% CI)	p	AUC (95% CI)	p
RISK6	70% (59–81)	<0.0001	70% (61–79)	0.0015	66% (46–86)	0.028
ACS 16-gene	67% (56–78)	0.0016	72% (63–81)	<0.0001	60% (41–79)	0.11
RISK4*	72% (55–88)	0.0054	72% (53–92)	0.0063	67% (50–83)	0.02

Table 1. Performance of RISK6 signature in the GC6 cohort by blinded validation, compared to the ACS 16-gene and RISK4 signatures. Samples from 0–24 months before TB diagnosis were included. *In the smaller GC6–74 test set only, because RISK4 was discovered in the GC6-74 training set (comprising Gambian and South African samples).

Cohort	RISK6 threshold at sensitivity set to $\geq 75\%$ ^a	Specificity at sensitivity set to $\geq 75\%$ ^a	RISK6 threshold at specificity set to $\geq 75\%$ ^b	Sensitivity at specificity set to $\geq 75\%$ ^b	Case samples	Control samples	AUC	AUC 95%CI
ACS (discovery)*	0.53	82.8%	0.46	82.8%	46	284	87.6%	82.8–92.4%
GC6–74*	0.35	50.0%	0.48	56.0%	50	372	70.6%	61.6–79.5%

Table 2. Accuracy of the RISK6 signature benchmarked against the WHO target product profile for prediction of incident TB within 12 months of sample collection. *Samples collected within 1 year of TB diagnosis. For GC6–74 participants from all three countries were combined. ^aSpecificity and the RISK6 threshold are reported at a sensitivity of 75%, which is the minimum criterion specified in the TPP for an incipient TB test⁴. At a sensitivity of 75%, the minimum specificity as set out in this TPP should be $\geq 75\%$. ^bSensitivity and the RISK6 threshold are reported at a specificity of 75%, which is the minimum criterion specified in the TPP for an incipient TB test⁴. At a specificity of 75%, the minimum sensitivity as set out in this TPP should be $\geq 75\%$.

	GC6-74 progressor vs non-progressor		
	AUC	95% CI	p
RISK6	70.2%	58.5–81.8%	0.0002
ACS 16-gene	67.6%	60.2–75.0%	<0.0001
Sweeney3	71.4%	61.2–81.7%	<0.0001
Maertzdorf4	65.7%	54.4–77.1%	0.0003

Table 3. Performance of RISK6 signature compared to the ACS 16-gene, Sweeney3 and Maertzdorf4 signatures in GC6-74 non-progressor vs progressor samples, collected within 6 months of TB diagnosis were included.

We also compared RISK6 to three other signatures (ACS 16-gene, Sweeney3⁹ and Maertzdorf4¹⁰, all measured by PCR), by comparing prognostic performance for discriminating between GC6-74 progressors and non-progressors on samples collected within 6 months of TB diagnosis (Table 3), as previously done by Warsinske *et al.* in the ACS cohort¹¹. Performance of the four signatures was equivalent, with AUC values ranging from 67.6% (ACS 16-gene, 95%CI 60.2–75.0) to 71.4% (Sweeney3, 95%CI 61.2–81.7). None of the signatures met the minimum criteria set out in the TPP for a progression test in the GC6-74 cohort (Table 3 and Supplementary Figure 3).

Because RISK6 was discovered on a South African cohort it was important to determine if performance varies by geography, since differences in population genetic structure, local epidemiology and environment may influence blood biomarker performance¹². We therefore also assessed prognostic performance by country. Interestingly, when assessing samples collected within 12 months of TB diagnosis, the AUC was highest for the Gambian cohort (AUC 76.3%, 95%CI 64.4–88.1%, Fig. 1f and Table 1), while the AUC for the South African cohort was similar to the entire, combined GC6-74 cohort (AUC 69.9%, 95%CI 55.5–84.2, Fig. 1f). Although the AUC for the smaller Ethiopian cohort (comprising 12 progressors) was also similar (AUC 69.3%, 95%CI 39.4–99.2), the confidence intervals were very large and discrimination between progressors and non-progressors was not significant (Fig. 1f).

Performance of RISK6 as a screening test in HIV-uninfected and HIV-infected individuals.

Expression levels of the six transcripts in RISK6 differed most between adolescent progressors and non-progressors at the time of TB diagnosis (Fig. 1a). In light of this, we hypothesized that RISK6 would also yield good performance as a screening or triage test for TB. Since HIV infection is a major risk factor for TB and a large proportion of TB patients in settings endemic for TB are HIV-infected¹³, we aimed to determine diagnostic performance in both HIV-infected and uninfected individuals. We therefore compared diagnostic performance of RISK6, benchmarked against the ACS 11-gene signature, in 112 HIV-uninfected (61 asymptomatic controls and 51 TB cases) and 82 HIV-infected (40 asymptomatic controls and 42 TB cases) adults from the Western Cape, South Africa. Excellent diagnostic performance of RISK6 was seen in both HIV-uninfected (AUC 93.7%, 95%CI 87.9–99.4%) and HIV-infected persons (AUC 92.6%, 95%CI 86.8–98.5); performance was not different between the two groups (DeLong unpaired test $p = 0.76$, Fig. 2a). By contrast, diagnostic performance

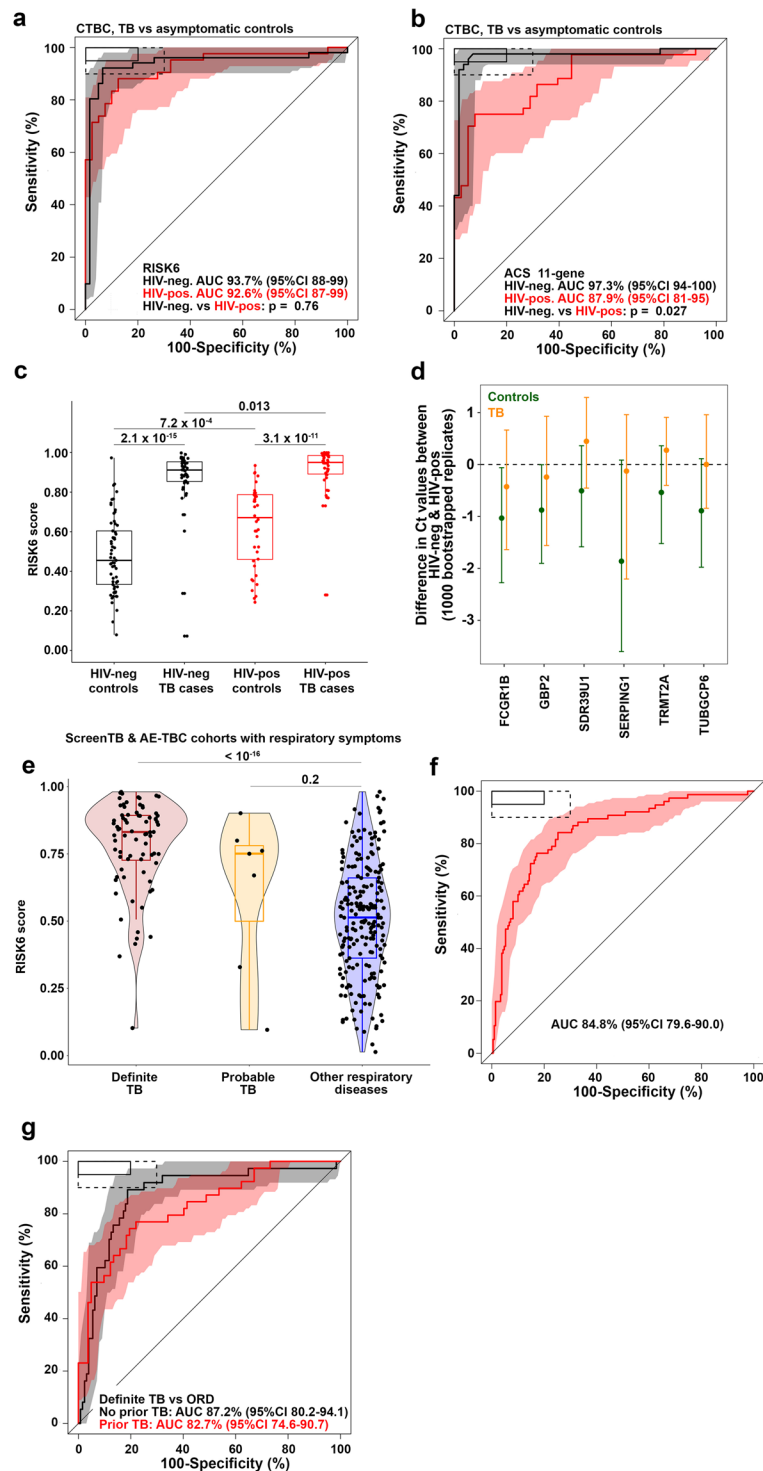


Figure 2. Diagnostic performance of RISK6 as a triage test. (a and b) ROC curves depicting diagnostic performance of RISK6 (a) and the ACS 11-gene signature (b), for discrimination between active TB cases and *Mtb*-infected controls in HIV-negative or HIV-positive individuals. RISK6 was measured by qRT-PCR on RNA from whole blood. Shaded areas depict the 95% CI. The boxes in the top left corner represent the optimal (solid line) and minimum (dotted line) criteria set out in the target product profile for a screening/triage test for TB. (c) Comparison of RISK6 signature scores in *Mtb*-infected controls and active TB cases in HIV-negative or HIV-positive individuals. P-values were calculated using the Mann-Whitney U test. Horizontal lines represent medians; boxes represent the IQR and whiskers the range. Dots represent individual sample scores. (d) Relative differences in RISK6 transcript expression levels between HIV-negative and HIV-positive *Mtb*-infected controls (green) or active TB cases (orange). Dots depict medians and error bars the 95% CI, calculated from 1000 bootstrapped Ct values. The dashed line represents zero. (e) Comparison of RISK6 signature scores, by blind prediction, in patients with definite TB (with rigorous microbiological confirmation), patients with probable

TB and in patients with other respiratory diseases (ORD). Horizontal lines depict medians, boxes the IQR and the whiskers the range. Violin plots depict the density of data points. P-values were computed by Mann-Whitney U test. (f) Receiver operating characteristic (ROC) curve depicting discrimination between diagnostic performance, by blind prediction of RISK6, measured by qRT-PCR in definite TB patients and patients with other respiratory diseases (ORD). Shaded areas depict the 95% CI. (g) ROC curves depicting RISK6 discrimination between definite TB and ORD participants of the ScreenTB and AE-TBC cohorts stratified into participants with no history of prior TB (black), or in those with a history of prior TB (red). Shaded areas depict the 95% CI.

of the ACS 11-gene signature was better in HIV-uninfected (AUC 97.3%, 95%CI 93.7–100) than in HIV-infected persons (AUC 87.9%, 95%CI 80.6–95.2); the 9% lower AUC in HIV-infected persons was significant (DeLong unpaired test $p = 0.027$, Fig. 2b). RISK6 signature scores were higher in HIV-infected controls compared to HIV-uninfected controls, suggesting an effect of underlying HIV infection on RISK6 (Fig. 2c). To understand the effects of underlying HIV infection on the RISK6 signature, we determined the difference in expression of each transcript between HIV-infected and uninfected individuals. Expression of FCGR1B and GBP2, but none of the other transcripts, was significantly higher in HIV-infected than uninfected controls, while no significant differences were observed in TB cases (Fig. 2d). At a sensitivity threshold of 90%, RISK6 achieved a specificity of 93.4% and 72.5% in HIV-uninfected and HIV-infected persons, respectively (Table 4 and Supplementary Table 2).

Diagnostic performance of RISK6 as a screening test in patients with respiratory symptoms.

We also determined RISK6 performance as a screening test in symptomatic adults enrolled into the ScreenTB¹⁴ and AE-TBC studies^{15,16}. These adult participants presented at primary health care clinics in Cape Town, South Africa with symptoms requiring investigation for TB including coughing for >2 weeks and at least another symptom consistent with TB, and were enrolled prior to the establishment of a TB or other disease diagnosis. RISK6 signature scores were measured on blinded PAXgene blood collected at presentation for care before treatment initiation in (1) 76 patients with microbiologically-confirmed, definite TB, (2) 7 patients with probable TB, and (3) 210 patients with other respiratory diseases (ORD) (Fig. 2e, see Supplementary Table 4 for diagnostic criteria). RISK6 discriminated between definite TB and ORD patients with an AUC of 84.8% (95%CI 79.6–90, Fig. 2f). At a sensitivity threshold of 90%, RISK6 achieved a specificity of 55.7% (95%CI 32.9–76.7) in these symptomatic patients (Table 4 and Supplementary Table 2), which falls short of the WHO target product profile for a community-based triage or referral test to identify people suspected of having TB⁵. RISK6 performance did not differ between HIV positive and negative participants (HIV-neg, $n = 250$: AUC 85.4%, 95%CI 79.7–91.0; HIV-pos, $n = 36$: AUC 79.5%, 95%CI 65.1–94.0, DeLong unpaired test $p = 0.46$, data not shown).

We unpacked the performance of RISK6 further in posthoc analyses performed after unblinding of patient diagnostic status. We noted that a considerable number ($n = 121$) of ScreenTB and AE-TBC participants had a record of least one previous episode of TB; this is typical of patients presenting for TB investigation in such high-incidence settings¹³. RISK6 discriminated between definite TB and ORD among participants with no history of prior TB with an AUC of 87.2% (95%CI 80.2–94.1%), whereas in those with a history of prior TB the AUC was 82.7% (95%CI 74.6–90.7%, Table 4, Fig. 2g). Although these AUCs were not statistically different, the specificities at a set sensitivity of >90% were markedly different at 75.0% and 37.8%, respectively (Table 4).

We also applied RISK6 to published microarray datasets, using RISK6geo, a risk score algorithm adapted for application to microarray or RNA-seq data, which is different to the RISK6 algorithm used throughout the manuscript when computing the RISK6 score from PCR data. RISK6geo discriminated between TB cases and asymptomatic *M.tb*-infected controls with AUCs exceeding 88% in all cohorts (Supplementary Table 3). To understand how these two score algorithms may differ, we compared performance characteristics of RISK6 and RISK6geo on qRT-PCR data from the different cohorts in this study. AUCs obtained with the two score algorithms were extremely similar and highly correlated (Spearman $\rho > 0.96$, Supplementary Table 2).

Performance of RISK6 as a TB treatment monitoring biomarker.

Diagnostic performance of RISK6 was also assessed in the Catalysis cohort of TB patients who were studied during and after TB treatment^{17–19}. RISK6 achieved an AUC of 93.5% (95%CI 85.5–100) for discriminating between newly diagnosed TB cases and asymptomatic controls (Fig. 3a). Next, we determined if RISK6 has utility as a biomarker for monitoring TB treatment. We hypothesized that RISK6 scores, which are very high in patients with active disease, would decrease rapidly during TB treatment such that samples collected after bacteriological cure can be discriminated from the respective pre-treatment sample with high accuracy. We also hypothesized that RISK6 would allow discrimination of cured patients from those with treatment failure after 24 weeks of treatment. When measured by qRT-PCR in patients with bacteriological cure in the Catalysis cohort, RISK6 scores decreased significantly during TB treatment, although scores observed at the end of treatment were still significantly higher than those observed in healthy controls (Fig. 3b), as reported for the 16-gene ACS signature previously¹⁹. Despite this, RISK6 significantly discriminated between samples collected pre-treatment and one week after treatment initiation (AUC 79.5%, 95%CI 72.2–86.7), four weeks after treatment initiation (AUC 77.4%, 95%CI 69.9–84.9) and end of treatment samples (AUC 88.1%, 95%CI 82.5–93.6; (Fig. 3c). Importantly, RISK6 was a strong predictor of treatment outcome and significantly differentiated between the 78 patients with bacteriological cure and the 7 patients with treatment failure even before treatment initiation (time of TB diagnosis, AUC 77.1, 95%CI 52.9–100, Fig. 3d), and at the end of treatment (AUC 95.2, 95%CI 87.5–100, Fig. 3d). These data are consistent with RISK6 detecting differences in inflammatory profiles before the initiation of treatment which predict the outcome of treatment, while also detecting ongoing inflammation in those who fail treatment and do not achieve bacteriological cure

Cohort (comparison)		RISK6 Threshold [‡]	Sensitivity, set to $\geq 90\%$	Specificity [‡]	Cases, n	Controls, n	AUC	AUC 95%CI
Cross-sectional TB (Definite TB vs asymptomatic controls)	HIV+	0.78	90.5%	72.5%	42	40	92.6%	86.8–98.5%
	HIV–	0.78	90.2%	93.4%	51	61	93.7%	87.9–99.4%
ScreenTB and AE-TBC: Symptomatic adults (Definite TB vs ORD)	A-priori analysis	0.55	90.8%	55.7%	76	210	84.8%	79.6–90.0%
	[§] Post-hoc analysis in those without previous TB	0.61	91.0%	75%	37	128	87.2%	80.2–94.1%
	[§] Post-hoc analysis in those with previous TB	0.46	92.3%	37.8%	39	82	82.7%	74.6–90.7%
RePORT-Brazil (Definite TB vs household contacts)	Definite TB vs combined QFT+ and QFT–	0.61	90.2%	73.7%	51	99	90.9%	85.2–96.6%
	Definite TB vs QFT+	0.61	90.2%	59.1%	51	22	88.6%	81.2–96%
	Definite TB vs QFT–	0.61	90.2%	77.9%	51	77	91.5%	85.9–97.1%
Peru (Definite TB vs household contacts)	Definite TB vs combined QFT+ and QFT–	0.21	91.7%	65.6%	48	96	90.6%	85.5–95.6%
	Definite TB vs QFT+	0.21	91.7%	57.1%	48	49	89.6%	83.5–95.7%
	Definite TB vs QFT–	0.19	91.7%	74.5%	48	47	91.5%	86.2–96.9%
Catalysis Cohort (Definite TB vs asymptomatic controls)		0.39	90.0%	95.2%	87	21	93.9%	85.5–100%

Table 4. Accuracy of the RISK6 signature benchmarked against the WHO target product profile for a screening/triage test. [‡]Specificities and the RISK6 threshold are reported at a sensitivity of $\sim 90\%$, which is the minimum criterion specified in the TPP for an incipient TB test⁴. At a sensitivity of $\sim 90\%$, the minimum specificity as set out in this TPP should be $\geq 70\%$. [§]Posthoc performance assessed after sample unblinding.

by 24 weeks. To address this further, we determined if blood RNA signature scores were associated with *in vivo* pulmonary inflammation measured by ¹⁸F-labeled fluorodeoxyglucose (¹⁸F FDG) PET-CT. Surprisingly, RISK6 scores directly correlated with metabolic activity in lung lesions as measured by total glycolytic activity index (TGAI) (Spearman $\rho = 0.66$, $p < 0.0001$, Fig. 3e), while signature scores correlated inversely with Xpert Ct values (Spearman $\rho = -0.60$, $p < 0.0001$) and Mycobacterial Growth Indicator Tube (MGIT) culture days-to-positivity values (Spearman $\rho = -0.67$, $p < 0.0001$) measured at TB diagnosis (data not shown).

Performance of RISK6 as triage test and TB treatment monitoring biomarker in South American Cohorts.

An important issue is how a biosignature that was trained and validated in African cohorts will perform in geographically distinct populations. To address this, we assessed diagnostic performance of RISK6 measured by qRT-PCR in cohorts from Peru and Brazil. In the Peruvian cohort, RISK6 discriminated between culture-positive TB patients and QFT-negative asymptomatic controls with an AUC of 91.5% (95%CI 86.2–96.9, Fig. 4a) and between TB patients and QFT-positive asymptomatic controls with an AUC of 89.6% (95%CI 83.5–95.7, Fig. 4b). RISK6 also achieved an AUC of 90.9% (95%CI 85.2–96.6) for discriminating between Brazilian culture-positive TB patients and close contacts (Fig. 4c). The minimum criteria for a screening or triage test for TB, set out by the WHO, set the sensitivity at $\geq 90\%$ at a specificity of 70%⁵. With sensitivity at $\geq 90\%$, performance of RISK6 in the South American cohorts met these criteria when discriminating between TB cases and QFT-negative controls, but discrimination between TB cases and QFT-positive controls fell short of the 70% specificity mark (Table 4 and Supplementary Table 2). It was notable that the RISK6 score threshold at which the sensitivity was $\sim 90\%$ was quite variable, suggesting that the positivity cut-off for a signature that predicts incident TB, detects prevalent TB, or monitors treatment response, may have to be different. The epidemiological setting and design features of each cohort, as well as experimental variability in the laboratory are likely to have contributed to the different thresholds we observed, highlighting the need for additional assay qualification. We note that the much lower threshold for the Peruvian cohort was because RISK6 was measured on RNA extracted from PBMC (not whole blood), which was previously shown to reduce transcriptomic signature scores, although discrimination between TB and controls was not affected⁷.

We also determined performance of RISK6 as a biosignature for monitoring TB treatment in the Brazilian patients, all of whom achieved microbiological cure after 6 months of treatment. RISK6 scores decreased significantly after 8 weeks of TB treatment and, unlike the Catalysis cohort, scores observed in the Brazilian patients at the end of treatment had reached levels observed in healthy controls (Fig. 4d,e). RISK6 also significantly discriminated between samples collected pre-treatment and 8 weeks after treatment initiation (AUC 69.2%, 95% CI 58.6–79.7), and end of treatment samples (AUC 87.3%, 95% CI 79.9–94.6, Fig. 4f).

RISK6 as a treatment biomarker in HIV-infected patients with recurrent TB.

The promising results from these treatment response studies prompted us to also evaluate if RISK6 can monitor success of recurrent TB treatment in HIV-infected individuals on ART, who participated in the randomized controlled IMPRESS trial. IMPRESS determined if a retreatment regimen that contained moxifloxacin, instead of ethambutol, would improve TB retreatment outcomes relative to the standard regimen²⁰. No differences in RISK6 scores were observed between the two treatment arms of the trial (data not shown). Consequently, all analyses were performed with the treatment arms combined. RISK6 scores decreased upon treatment (Fig. 5a) and could discriminate significantly between samples collected at the pre-treatment time point (baseline) and those collected after the intensive phase of treatment, at 2 months (AUC 75.1%, 95%CI 66.5–83.8, Fig. 5b). Discrimination between

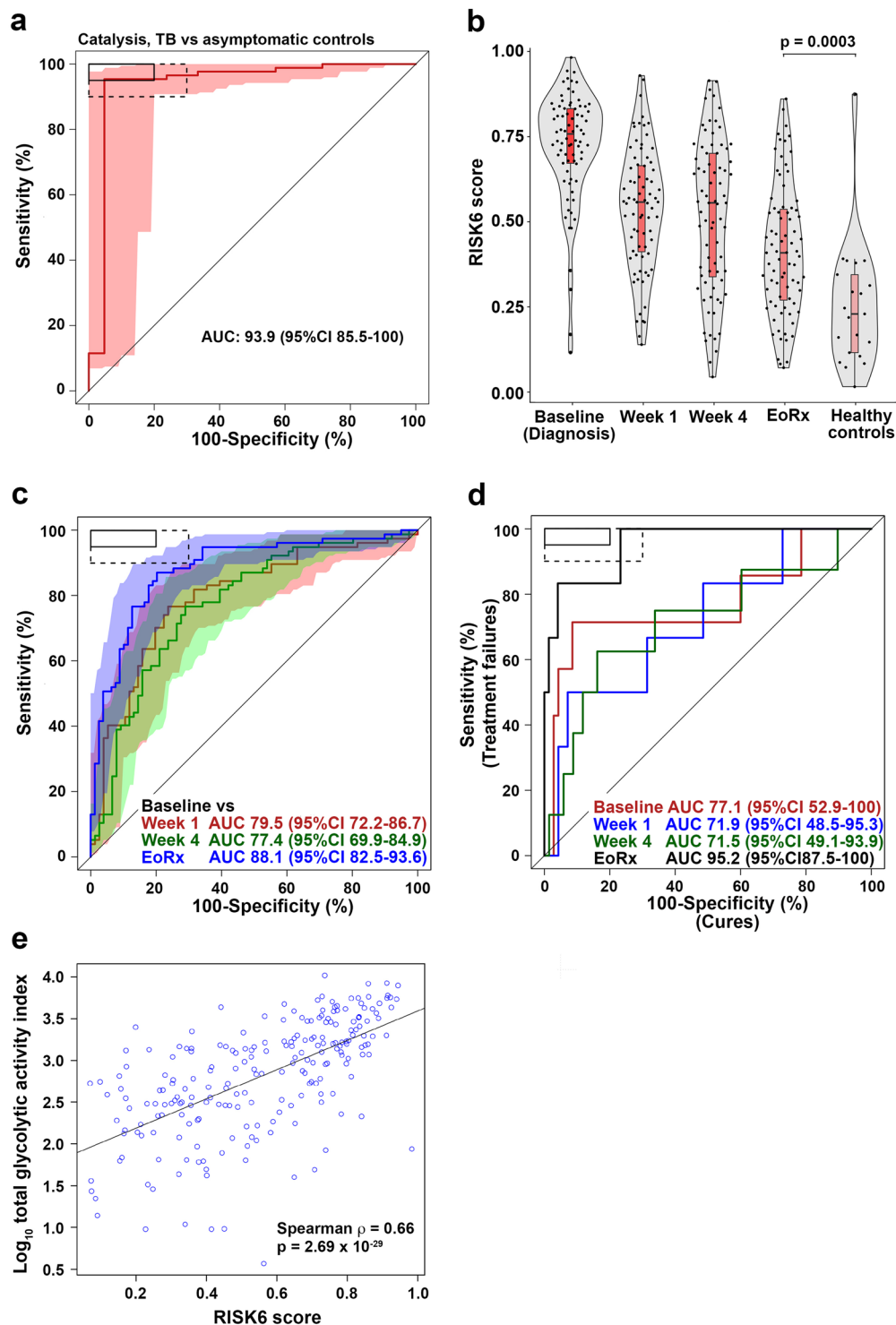


Figure 3. Treatment monitoring using the RISK6 signature in the Catalysis TB Treatment Cohort. **(a)** ROC curve depicting diagnostic performance of RISK6 for discriminating between active TB cases (irrespective of treatment outcome), sampled prior to treatment initiation (baseline, $n = 87$), and controls ($n = 21$) from the Catalysis Cohort. Shaded areas depict the 95% CI. The boxes in the top left corner represent the optimal (solid line) and minimum (dotted line) criteria set out in the target product profile for a screening/triage test for TB. **(b)** Comparison of RISK6 signature scores in cases (irrespective of treatment outcome) from the Catalysis Cohort at baseline and week 1 or week 4 after treatment initiation and after treatment completion (EoRx). Also shown are the RISK6 signature scores in healthy controls. Horizontal lines depict medians, the boxes the IQR and the whiskers the range. Violin plots depict the density of data points. The p -value, computed by Mann-Whitney U test, compares RISK6 signature scores after treatment completion with those in controls. **(c)** ROC curves depicting performance of RISK6 for discriminating between baseline (pre-treatment) samples and samples collected after week 1, week 4 or completion of TB treatment. **(d)** Prediction of treatment failure using

RISK6. ROC curves depict discrimination between cases with cure ($n = 70$) and those with treatment failure ($n = 7$) in samples collected at treatment initiation (Baseline, red), at week 1 (blue), week 4 (green) and after treatment completion (EoRx, black). (e) RISK6 scores plotted versus total glycolytic activity index measured by PET-CT for all available samples.

baseline samples and those collected at the end of treatment, when all patients had achieved clinical cure, was better than after 2 months of treatment (AUC 91.2%, 95%CI 86.0–96.3, Fig. 5b), although inflammation appeared to resolve further after the end of treatment, since RISK6 discriminated best between baseline and samples collected 6–8 months after treatment completion (AUC 98.5%, 95%CI 96.5–100, Fig. 5b). When measured at baseline (not shown) or the end of the intensive treatment phase at 2 months (AUC 63.4%, 95%CI 48.2–78.5), RISK6 did not discriminate significantly between patients who had sputum culture conversion at 2 months and those who converted after 2 months (Fig. 5c). It was not possible to determine if RISK6 could predict treatment failure in this trial since all patients achieved bacteriological cure.

Although HIV infection causes immunodeficiency, it also drives chronic immune activation and inflammation^{21–23} and induces expression of type I IFN response, including interferon stimulated genes (ISGs)^{24,25}. Successful antiretroviral therapy (ART) suppresses viral replication and reduces plasma viral load (pVL), decreasing inflammation and immune activation, although not to levels typical of HIV-uninfected persons²⁶. Since RISK6 includes three IFN-inducible ISG transcripts, we aimed to evaluate the effect of pVL on signature scores in the IMPRESS trial. Eighty-five participants had pVL measurements, 36 with detectable viral loads (above 400 copies per mL) and 49 with undetectable viral loads (below 400 copies per mL); sixty of the measurements were baseline samples and 25 were end-of-treatment samples. RISK6 signature scores were significantly higher in samples with detectable pVL than those with undetectable pVL ($p = 0.0027$, Fig. 5d), showing that pVL is a confounder in ISG-containing transcriptomic signatures.

Robustness of the PCR-based RISK6 signature. An advantage of the pair-wise ensemble structure of RISK6 is that a signature score can be calculated even if one or more transcript is not detected, for example due to a failure during PCR amplification. To determine how robust the signature is to such missing data, we compared diagnostic performance for discriminating between HIV-uninfected TB cases and asymptomatic controls (Fig. 2) by the full 6-gene RISK6 signature, which comprises nine pairs formed between six transcripts, or after removing one, two, three or four, of these transcripts such that every combination of the pairs was tested. Diagnostic performance was not affected by removal of a single transcript, irrespective of transcript identity (AUC for full RISK6: 93.6%, 95%CI 87.4–99.7; average AUC for 5-transcript signature: 93.2%, lower 95%CI bound: 85.5, Fig. 6a). However, removal of two or more transcripts, especially when two or more of SERPING1, SDR39U1 or TUBGCP6 were omitted, resulted in somewhat decreased performance of RISK6 (average AUC for 4-transcript signature: 92.4%, lower 95%CI bound: 80.1; average AUC for 3-transcript signature: 91.4%, lower 95%CI bound: 72.1). A very similar result was observed when the same analysis was performed on the Brazilian cohort (Supplementary Figure 4). These results show that RISK6 can tolerate one or even two missing transcripts without the diagnostic performance being markedly eroded.

Effective deployment of transcriptomic signature tests such as RISK6 in community or primary health care settings is dependent on successful translation of gene expression quantification to methods that are simple, cheap and rapid. An expensive and cumbersome component of any blood transcriptomic assay is the procedure and cost of blood collection. Therefore, we sought to determine if RISK6 could be reliably measured in very small volumes of capillary blood collected by finger stick. We compared discrimination between healthy controls and TB cases by RISK6, measured by qRT-PCR in 20 μ L, 50 μ L or 100 μ L capillary blood, benchmarked against the typical 2.5 mL venous blood collected in PAXgene tubes. Among samples collected from the 49 participants, the number of samples with one or more failed PCR reaction, where the amplification curve for one transcript did not pass the QC threshold defined by Fluidigm, for the 20 μ L, 50 μ L and 100 μ L capillary blood volumes was 4 (8%), 3 (6%) and 3 (6%), respectively. None of the 2.5 mL venous blood samples yielded failed PCR reactions. When failure of 1 of the 6 transcripts was tolerated, RISK6 scores could be calculated for 98% (1 failed sample), 98% and 100% of the 20 μ L, 50 μ L and 100 μ L capillary blood samples, respectively.

RISK6 signature scores measured on 2.5 mL venous blood correlated strongly with those measured on 20 μ L, 50 μ L or 100 μ L capillary blood samples (Spearman $\rho > 0.83$; Fig. 6b–d). Diagnostic performance of RISK6 was statistically non-inferior when measured in 20 μ L, 50 μ L or 100 μ L capillary blood samples compared with venous blood; ROC analysis yielded equivalent AUC curves (Fig. 6b–d). These results show that RISK6 can be measured on very small volumes of capillary blood collected by finger stick, which may be amenable to translation to a point of care testing platform.

Discussion

We discovered and validated RISK6, a parsimonious and robust blood transcriptomic signature with applicability for predicting incident TB, as a triage test for identifying individuals with or without respiratory symptoms who should be further investigated for TB disease, and for monitoring the response to TB treatment.

RISK6 identified individuals at risk of progression to incident TB and statistically validated in the independent GC6-74 cohort of TB household contacts by blind prediction. When applied to samples collected within 1 year of TB diagnosis in the ACS discovery cohort, RISK6 met the minimum criteria for a test for progression to TB set out by the WHO and FIND⁴. However, when applied to samples collected within 1 year of TB diagnosis in the GC6-74 validation cohort, the specificity at a sensitivity of $\geq 75\%$ was 50.3%, which did not meet these criteria (Table 2). Prognostic performance for incident TB of RISK6 was significantly better than that reported for the

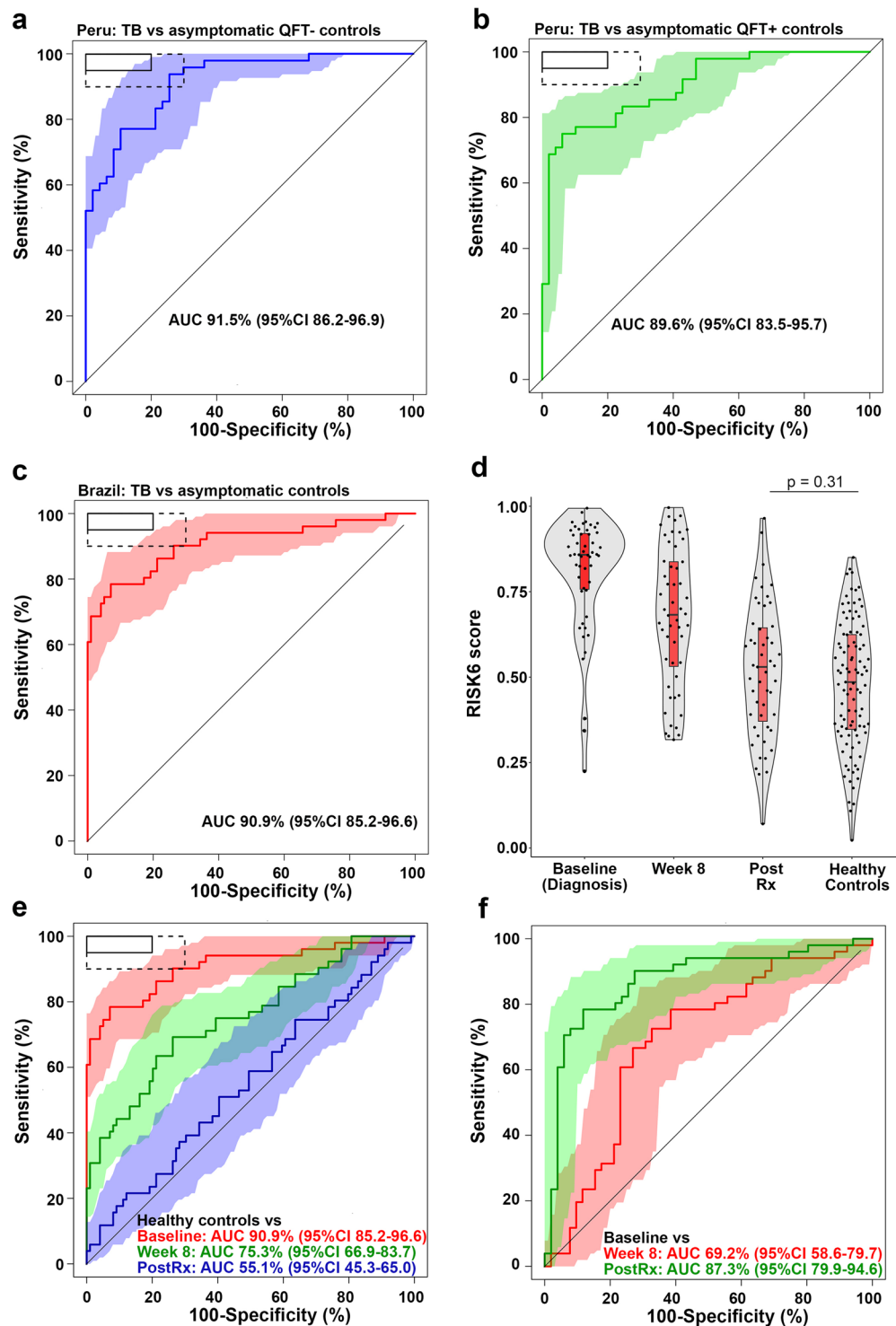


Figure 4. Diagnostic performance and treatment monitoring in South American cohorts. (a-b) ROC curve depicting diagnostic performance of RISK6 for discriminating between (a) culture-positive active TB cases ($n = 48$) and QuantiFERON-negative controls ($n = 47$), or between (b) culture-positive active TB cases ($n = 48$) and QuantiFERON-positive controls ($n = 49$) from the Peru Cohort. Shaded areas depict the 95% CI. The boxes in the top left corner represent the optimal (solid line) and minimum (dotted line) criteria set out in the target product profile for a screening/triage test for TB. (c) ROC curve depicting diagnostic performance of RISK6 for discriminating between culture-positive active TB cases, sampled prior to treatment initiation (baseline, $n = 51$), and controls ($n = 99$) from the RePORT-Brazil Cohort. (d) Comparison of RISK6 signature scores in TB cases at baseline, week 8 after treatment initiation and after treatment completion (Post Rx). Also shown are the RISK6 signature scores in healthy controls from Brazil. Horizontal lines depict medians, the boxes the IQR and the whiskers the range. Violin plots depict the density of data points. The p -value, computed by Mann-Whitney U test, compares RISK6 signature scores after treatment completion with those in controls. (e)

ROC curves depicting performance of RISK6 for discriminating between healthy control samples and samples collected from TB cases before treatment initiation (baseline), at week 8 after treatment initiation, or completion of TB treatment (Post Rx). (f) ROC curves depicting performance of RISK6 for discriminating between baseline samples from TB cases and samples collected 8 weeks after treatment initiation, or upon completion of TB treatment (Post Rx).

previously described 16-gene ACS signature⁶, which was discovered by RNA-seq also in the ACS progressor and non-progressor cohort. Ability to discriminate between GC6-74 non-progressors and progressors within 6 months of TB diagnosis was not different between RISK6 and the ACS 16-gene, Sweeney3 and Maertzdorf4 signatures, suggesting that these signatures detect similar biological processes that occur during disease progression. None of these signatures met the minimum TPP criteria for a progression test in the GC6-74⁴. The marked reduction in signature performance between ACS and GC6-74 was a notable result. Since these four signatures showed very similar performance in the GC6-74 cohort, we posit that it is most likely due to differences in age, epidemiology, study design and other factors between the cohorts. We also cannot exclude the possibility that reinfection with *M.tb* in some of the ACS participants, who reside in a setting of a very high force of infection²⁷, may have contributed to rapid disease progression, leading to a stronger transcriptomic signal.

Our findings are consistent with recent work from Gupta *et al.*, who systematically compared the prognostic performance of 17 published host-derived transcriptional signatures in a RNA-seq dataset that combined 4 progressor and controls cohorts, which included the ACS and GC6-74 cohorts²⁸. Despite the heterogeneity of cohorts and study designs used to derive these signatures, eight signatures including ACS 16-gene, Sweeney3 and RISK4 performed equivalently. Similarly, Warsinske *et al.* reported a comparison of 16 signatures for diagnostic performance in 24 datasets from TB cases and controls¹¹. They found that two signatures (Sweeney3⁹ and Sambarey10²⁹) met the TPP criteria for a triage test in data restricted to patients with culture-confirmed diagnosis of TB. Such head-to-head comparisons of different transcriptomic signatures are extremely valuable to allow unbiased selection of the best performing signatures for further evaluation.

Performance of RISK6 in the distinct cohorts from 3 different African countries was similar, although RISK6 did not significantly discriminate between progressors and non-progressors from Ethiopia, likely due to the small number of progressors, namely 12. The limitations of such small sample sizes for biomarker validation is also evident from other biomarker studies on the GC6-74 cohort and it was notable that the performance of RISK6 in the three GC6-74 cohorts was very similar to the previously published RISK4 signature (Table 2), which was specifically developed as a “pan-African signature”⁸. Recent studies showed that the 16-gene ACS signature, as well as the other small diagnostic Sweeney3 and Maertzdorf4 signatures, did not validate in either one or both of the GC6-74 validation sub-cohorts of Gambian or Ethiopian progressor and non-progressor TB household contacts⁸. However, when the 16-gene ACS signature was measured in the full GC6-74 cohort from The Gambia, comprising 30 progressors and 129 non-progressors⁶, the signature significantly validated by blind prediction. These results highlight the value of longitudinal cohort studies with sufficient incident TB cases to allow reliable assessment of prognostic performance of risk signatures. It is critical that more such cohort studies be performed to increase our collective capacity to develop, refine and validate such biomarkers.

A reliable and simple triage test to identify those who should be investigated more intensively for subclinical or active TB disease is urgently needed to improve case finding strategies and allow earlier diagnosis and treatment. RISK6 also performed well as a triage test in patients with respiratory symptoms who presented for care. However, with 56% specificity at >90% sensitivity, it did not meet the minimum criteria set out in the WHO target product profile (TPP) for a referral test to identify people who may have TB⁵. However, in our post-hoc analyses the specificity of RISK6 in differentiating between definite TB cases and ORD among patients with no prior history of TB was 75% at a set sensitivity of >90% (Table 4), which met the WHO target product profile (TPP). Data regarding the interval since the previous TB episode was often unavailable, precluding analysis of this factor. This finding highlights the importance of including clinical and epidemiological factors in studies of diagnostic biosignatures. Community-based case finding studies and prevalence surveys have shown that a substantial proportion of microbiologically-confirmed TB cases are asymptomatic³⁰⁻³², highlighting the need for TB case finding in asymptomatic communities. Definitive diagnoses of ORD patients were not determined in the ScreenTB and AE-TBC study, but larger future studies with careful diagnoses of patients with ORD would be essential to evaluate the specificity of RISK6, and define respiratory diseases most difficult to differentiate from TB.

RISK6 showed excellent diagnostic performance in differentiating between symptomatic TB cases and asymptomatic controls in four different case-control cohorts from South Africa, Peru and Brazil. Application of RISK6 to the South African cohorts met or exceeded the sensitivity and specificity criteria set out in the TPP for a screening or triage test⁵. In the South American cohorts, however, these criteria were only met when TB cases were compared to uninfected controls as determined by negative QuantiFERON tests. Whether this reflects a real geographic, genetic, environmental or epidemiological difference between South African and South American communities is not clear. For the Peruvian cohort RISK6 measurements were performed on RNA isolated from PBMC, which may have affected diagnostic performance, although we showed that near-identical ROC AUC results were observed when diagnostic performance of the ACS 11-gene signature was measured in whole blood and PBMC⁷. It is noteworthy that diagnostic performance of RISK6 was higher in Brazilian culture + smear + TB cases (AUC 99.8%, 95%CI 99.4–100) than in culture+smear- TB cases (AUC 90.5%, 95%CI 76.8–100) and that RISK6 scores correlated significantly with lung lesion activity measured by PET in the South African Catalysis cohort. RISK6 scores also decreased during disease resolution upon TB treatment and showed promise as a treatment response biomarker. This reflects the opposite of the increasing inflammatory signals detected by RISK6 during disease progression, as previously reported for other transcriptomic signatures¹⁹. Our findings strongly

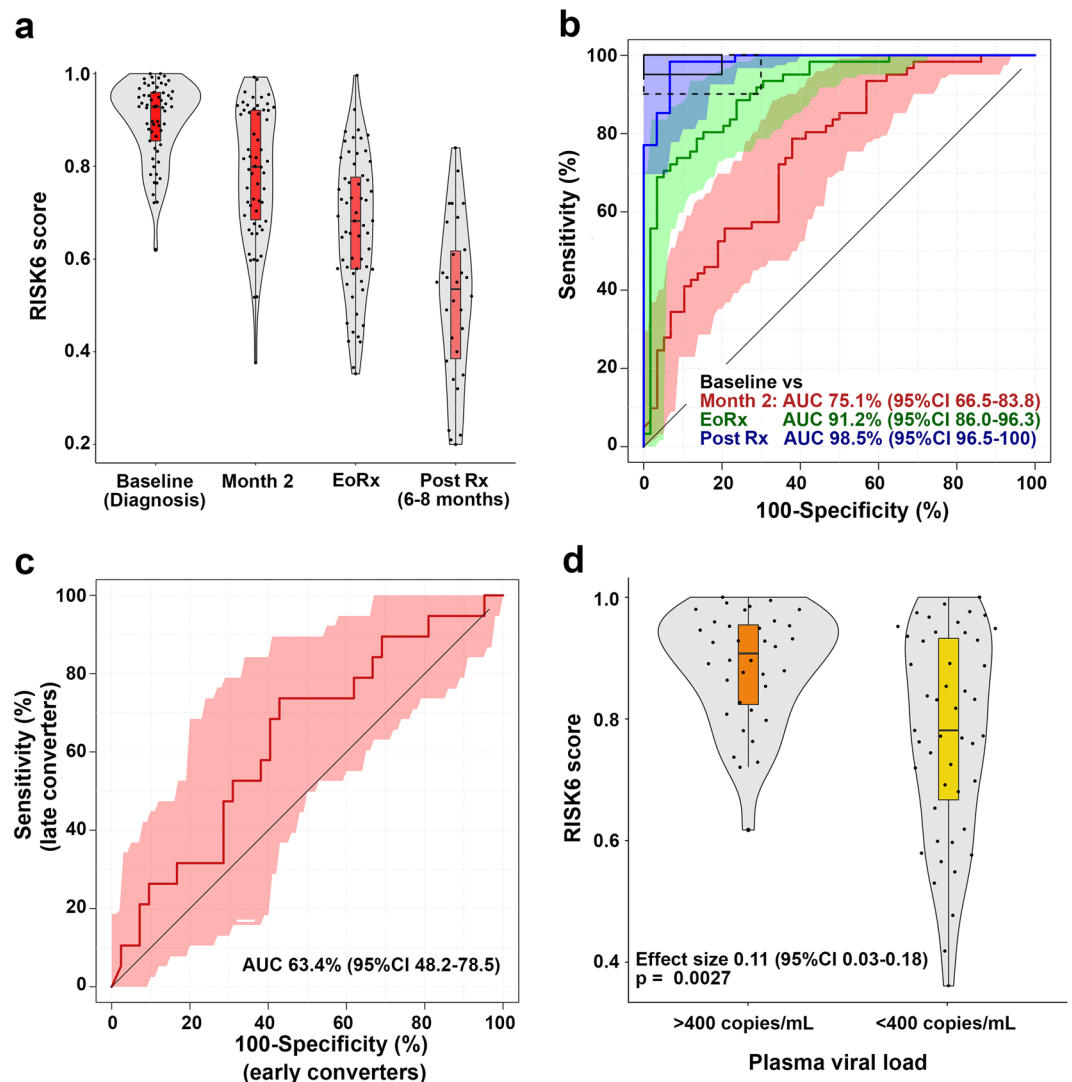


Figure 5. Treatment monitoring using the RISK6 signature in the IMPRESS cohort. **(a)** Comparison of RISK6 signature scores in HIV-infected TB cases (irrespective of treatment outcome) from the IMPRESS cohort at baseline, month 2 after treatment initiation, at treatment completion (EoRx) and 6–8 months after treatment completion (Post Rx). Horizontal lines depict medians, the boxes the IQR and the whiskers the range. Violin plots depict the density of data points. **(b)** ROC curves depicting performance of RISK6 for discriminating between baseline (pre-treatment) samples and samples collected at month 2 after treatment initiation, at treatment completion (EoRx) and 6–8 months after treatment completion (Post Rx). Shaded areas depict the 95% CI. The boxes in the top left corner represent the optimal (solid line) and minimum (dotted line) criteria set out in the target product profile for a screening/triage test for TB. **(c)** ROC curve depicting performance of RISK6, measured at 2 months on TB treatment, for discriminating between TB cases who converted their sputum to negative by 2 months (early converters) and those who converted after 2 months. **(d)** Comparison of RISK6 signature scores in IMPRESS participants stratified by detectable (>400 RNA copies/mL plasma) vs. undetectable plasma HIV load (<400 RNA copies/mL plasma). Horizontal lines depict medians, the boxes the IQR and the whiskers the range. Violin plots depict the density of data points. The effect size is the relative difference in median plasma viral load and the p-value was calculated using Mann-Whitney U test.

suggest that disease severity in TB cases plays a role in performance of transcriptomic signatures, as reported previously for other signatures^{19,33}, and likely other biomarkers. Given the lines of evidence that such signatures track severity of disease and lung lesions, it should be noted that biomarker performance in populations from different settings may be influenced by differences in study design that may preferentially enrol patients with more or less severe disease, rather than reflecting purely geography-associated differences. Larger and well-designed longitudinal biomarker studies are necessary to investigate the performance characteristics of blood biomarkers, such as RISK6, for classifying individuals with ambiguous respiratory phenotypes that are difficult to diagnose, and for revealing which stage of the TB spectrum such individuals may fall into.

Underlying HIV-infection did not significantly affect diagnostic or treatment response performance of RISK6, which is crucial given the high prevalence of undiagnosed TB in people living with HIV¹³. We acknowledge

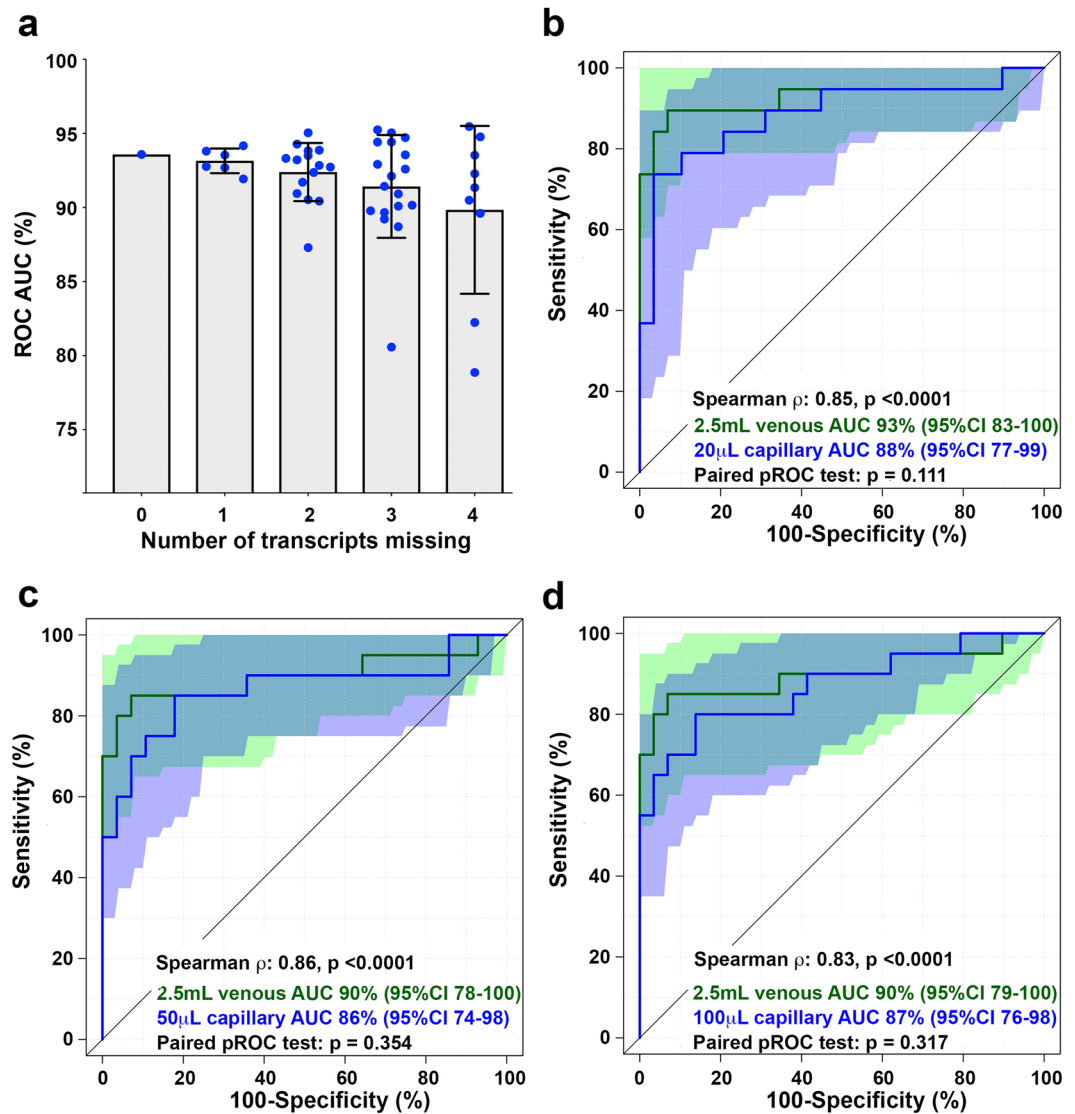


Figure 6. Robustness of RISK6. **(a)** ROC AUC values for discrimination between HIV-uninfected TB cases and asymptomatic controls in the CTBC cohort by the full 6-gene RISK6 signature (9 pairs formed between 6 transcripts, far left), or after removing 1, 2, 3 or 4 of the transcripts such that every combination of the pairs (represented by individual blue dots) was tested. The grey bar graph represents the mean and the error bar the 95% CI. **(b–d)** ROC curves depicting performance of RISK6 for discriminating between TB cases and controls in the capillary blood cohort. Green curves represent RISK6 measured on 2.5 mL venous blood collected in PAXgene tubes. Blue curves represent RISK6 measured on 20 μ L **(b)**, 50 μ L **(c)** or 100 μ L **(d)** capillary blood collected by finger stick. Shaded areas depict the 95% CI. The Spearman correlation coefficients and associated p-values for comparison of RISK6 scores measured by 2.5 mL venous blood versus each capillary blood volume are shown. Only individuals with paired venous and capillary blood results were included in each comparison (i.e. venous blood results in those with a missing capillary blood value were excluded from each graph). ROC curves were compared using the DeLong paired test in R and resulting p-values are shown.

that the effect of HIV was not assessed in all of the validation cohorts and more such analyses are necessary to definitively establish the effects of underlying HIV infection on RISK6 performance. Regardless, other published blood-based transcriptomic TB signatures showed reduced diagnostic performance in HIV-infected compared to uninfected persons^{6,9,26–28}. Since most transcriptomic TB signatures detect the elevation of ISG expression during TB, this effect of HIV is not surprising given that strong Type I IFN responses constitute the typical anti-viral response³⁴. Persistent HIV viremia also drives chronic immune activation³⁵ which is characterised by high ISG expression. Our data show that HIV-infection was associated with elevated RISK6 signature scores, but that expression levels of individual transcripts in the signature were not dramatically modulated by HIV-infection. Although discrimination between HIV-infected TB cases and controls was not diminished relative to HIV-uninfected people, our results show that a different diagnostic test threshold would likely be required for HIV-uninfected and HIV-infected populations. A limitation of our analyses of HIV effects is that the clinical studies were not sufficiently powered to investigate the performance of RISK6 in samples with detectable or high

pVLs. The issue of limited power will be addressed in a prospective, multicohort study currently underway in South Africa by performing a head-to-head comparison of performance of RISK6 with other signatures, in both HIV-uninfected and HIV-infected persons (clinicaltrials.gov NCT02735590). Our work suggests that underlying HIV infection has a marked effect on performance of IFN response signatures, which requires further examination. Of note, Esmail, Wilkinson and colleagues demonstrated that a transcriptomic TB signature based on complement pathway genes may have greater utility in ART naïve HIV-infected persons³⁶. In their study, pVL did not affect circulating immune complexes, which were associated with transcripts involved in the complement pathway.

We found that RISK6 scores correlated significantly with lung lesion activity measured by PET-CT in TB patients of the Catalysis study who underwent TB treatment. RISK6 showed good performance as a treatment response biomarker, decreasing in score during successful treatment and showing very good discrimination between pre-treatment and post-treatment samples in patients with clinical cure, even in patients with underlying HIV-infection. Similar utility as a treatment response biomarker was observed in the Brazilian cohort and the South African Catalysis cohort. Importantly, in the Catalysis cohort RISK6 significantly predicted treatment failure prior to treatment initiation and differentiated between treatment failures and cured patients with very high accuracy at the end of treatment. These findings suggest that RISK6 detects inflammatory signals associated with the TB disease process in the lungs or other affected sites and that resolution of these processes can be tracked by monitoring gene expression in the blood. Our data provide proof of concept that RISK6 allows treatment monitoring, as has been shown for a number of other transcriptomic signatures^{19,33,37,38}.

We explicitly developed RISK6 with the ultimate objective of translation to a hand-held point-of-care platform and therefore conducted all performance analyses, including the training and all validation cohorts, by qRT-PCR using a standardized protocol and locked-down analysis algorithm. The RISK6 score is computed based on an ensemble of nine transcript pairs using the pair-ratio approach, which uses ratios of transcripts regulated in opposite directions during TB progression, as previously described^{8,11,36,37}. This pair-ratio feature of RISK6 eliminates the need for standardisation of gene expression using reference (or housekeeper) transcripts, restricting measurement of the signature to six primer-probes and simplifying data processing steps. Importantly, RISK6 was measured by a highly standardized, locked-down protocol in our studies. Consequently, RISK6 performance was not subject to the gene expression normalization methods that are typically necessary to overcome reproducibility problems due to sample and batch effects associated with microarray and RNA-sequencing data^{39,40}. Regardless, to allow measurement of RISK6 scores in public microarray or RNA-sequencing datasets, we also provide the score computation algorithm, “RISK6geo”, which computes virtually equivalent scores to the RISK6 algorithm from qRT-PCR data. Finally, we showed that RISK6 could be measured on very small volumes of capillary blood collected by fingerstick, with no discernible effect on signature performance.

Our results support work towards incorporation of RISK6 into rapid, capillary-blood-based point-of-care devices for field evaluation in community and primary care settings and implementation studies.

Methods

We developed the transcriptomic signature of risk, RISK6, using samples collected from participants of the Adolescent Cohort Study (Supplementary Figure 1a). RISK6 was then applied to seven external validation cohorts to determine prognostic and diagnostic performance and utility as a treatment response biomarker (Supplementary Figure 1b). Most of these cohorts have been described previously^{6,8,19,41,42}.

Adolescent cohort study (ACS) (RISK6 discovery). The Adolescent Cohort Study, including selection of progressors and non-progressors, was previously described^{6,8,41}. Briefly, among 6,363 healthy adolescents from the Worcester region of the Western Cape, South Africa, who were enrolled, 46 “progressors” were either TST or QuantiFERON TB-Gold In-Tube assay (Qiagen) (QFT)-positive and developed microbiologically-confirmed intrathoracic disease during 2 years of follow-up. Individuals who were TST or QFT-positive at enrolment and remained healthy (no TB disease) during follow-up, and matched the progressors for age, gender, ethnicity, school of attendance and prior history of TB disease, were included as “non-progressors”. Participants were excluded if they developed tuberculosis disease within 6 months of enrolment (or the first TST or IGRA-positive sample) to exclude early asymptomatic disease that could have been present at the time of assessment, or if they were HIV-infected. Longitudinally collected PAXgene samples were available from most participants at six-monthly intervals. The Human Research Ethics Committee of the University of Cape Town approved the study (045/2005) and all participants provided written, informed assent, while parents or legal guardians provided written, informed consent. All research was performed in accordance with relevant guidelines/regulations.

GC6-74 cohort (prognostic validation). The Grand Challenges 6–74 project was previously described^{6,8,43}. Briefly, HIV-uninfected household contacts of TB cases were longitudinally followed for up to 2 years, with assessments at baseline, at 6 months and at 18 months. TB progressors who developed microbiologically confirmed TB during follow-up were retrospectively identified and matched 1:4 to healthy non-progressors. Individuals in whom TB disease developed within 3 months of baseline were excluded. PAXgene samples collected from 26 Gambian progressors and 116 non-progressors, 41 South African progressors and 164 non-progressors, and 12 Ethiopian progressors and 48 non-progressors were included. Participants provided written, informed consent. Protocols were approved by the Joint Medical Research Council and Gambian Government Ethics Review Committee, Banjul, The Gambia (SCC.1141vs2), the Stellenbosch University Institutional Review Board (N05/11/187) and the Armauer Hansen Research Institute (AHRI) / All Africa Leprosy, TB and Rehabilitation Training Center (ALERT) Ethics Review Committee (P015/10). All research was performed in accordance with relevant guidelines/regulations.

Cross-sectional TB cohort (CTBC, Diagnostic validation). Adults with newly diagnosed active TB (sputum Xpert MTB/RIF-positive or liquid culture-positive) were recruited at primary healthcare clinics in Worcester and Masiphumelele, South Africa. Asymptomatic community controls were recruited from the Worcester or Masiphumelele areas. HIV-infection was diagnosed with the Determine HIV1/2 test (Alere). Protocols were reviewed and approved by the Human Research Ethics Committee of the Faculty of Health Sciences at the University of Cape Town (HREC 126/2006 and HREC 288/2008). All study participants provided written informed consent and all research was performed in accordance with relevant guidelines/regulations.

Participants included 112 HIV-uninfected adults (51 TB cases and 61 asymptomatic controls) and 82 HIV-infected (42 TB cases and 40 asymptomatic controls). Blood was collected in PAXgene tubes at diagnosis in TB cases and at enrolment in asymptomatic controls.

ScreenTB and AE-TBC cohorts (Diagnostic validation). Adults aged >18 years who presented at primary health care clinics in Cape Town or emergency or medical wards of Tygerberg Hospital in Cape Town with respiratory symptoms compatible with TB, including cough for at least 2 weeks and another symptom including fever, weight loss, haemoptysis or night sweats, were screened for inclusion for the ScreenTB¹⁴ or the African-European Tuberculosis Consortium (AE-TBC) studies^{15,16}. Those who had TB treatment within 90 days, received immunosuppressive medication (ScreenTB) or quinolones or aminoglycosides in the past 60 days (AE-TBC), had a record of alcohol or drug abuse or a haemoglobin level <9 g/dL (ScreenTB) or <10 g/dL (AE-TBC), or who were pregnant or breastfeeding, where not eligible. HIV-infection was not an exclusion criterion. Using a pre-defined TB classification algorithm¹⁵ (Supplementary Table 4), patients with microbiologically confirmed pulmonary TB were classified as having definite TB (n = 76). Those with either a single positive sputum smear or with chest radiographs that were compatible with pulmonary TB and who responded to TB treatment were classified as probable TB (n = 7). Patients whose sputum tested negative, and who were not started on TB treatment were classified as having “other respiratory diseases” (ORD, n = 210). These ORD patients also did not have a TB diagnosis during 2 months further follow-up. We also performed a post-hoc analysis after unblinding, where study participants were grouped into those with a previous history of TB (TB cases, n = 39; ORD, n = 82) and those without a previous history of TB (TB cases, n = 37; ORD, n = 128). All study participants provided written, informed consent. The documents for the ScreenTB and AE-TBC studies were approved by the Health Research Ethics Committee at Stellenbosch University and all research was performed in accordance with relevant guidelines/regulations. Blood was collected in PAXgene tubes at enrolment, before treatment initiation.

Peruvian household contacts cohort (Diagnostic validation). Bacillus Calmette-Guérin (BCG)-vaccinated, HIV-uninfected Peruvian participants were recruited through Socios En Salud (SES), an affiliate of Partners in Health from urban and peri-urban settlements around Lima, Peru, as a case-control study. Participants included adults with recently diagnosed microbiologically confirmed, culture-positive, drug-sensitive pulmonary TB disease (active TB, n = 48), and clinically asymptomatic household contacts of TB patients assessed within two-weeks of diagnosing the index case. Household contacts were evaluated for signs of TB disease at the time of enrolment, and were excluded if clinical symptoms of TB were present. Healthy household contacts were assessed for *M.tb* infection using QuantiFERON TB-Gold In-Tube (QFT) assays. Participants with QFT IFN γ responses ≥ 0.35 international units (IU)/mL were considered latently *M.tb* infected (QFT-positive, n = 49) and uninfected if QFT IFN γ < 0.35 IU/mL (QFT-negative, n = 47). Household contacts were evaluated for signs of TB disease at the time of enrolment, and were excluded when clinical symptoms of TB were present. The Institutional Review Board of the Harvard Faculty of Medicine and Partners Healthcare (protocol number IRB16-1173), and the Institutional Committee of Ethics in Research of the Peruvian Institutes of Health approved the study protocol. All adult study participants and parents and/or legal guardians of minors provided informed consent, while minors provided assent. All research was performed in accordance with relevant guidelines/regulations. Peripheral blood mononuclear cells (PBMC) were isolated from 50 mL of venous blood using ficoll and cryopreserved at 5×10^6 cells/cryovial, then shipped to the Brigham and Women’s Hospital for storage. RNA was extracted from 10^6 cells PBMCs using the RNeasy extraction kit (Qiagen).

RePORT-Brazil cohort (Diagnostic validation and treatment response). Regional Prospective Observational Research for Tuberculosis (RePORT)-Brazil is an ongoing prospective cohort study at five participating centers in Brazil: three in Rio de Janeiro (Instituto Nacional de Infectologia (INI), Clinica de Saude Rinaldo Delmare (Rochina), Secretaria de Saude de Duque de Caxias (Caxias), one in Salvador (Instituto Brasileiro para Investigaçao da Tuberculose), and one in Manaus (Fundação Medicina Tropical Dr. Heitor Vieira Dourado). RePORT-Brazil enrolls participants ≥ 18 years-old who initiate treatment for culture-confirmed pulmonary TB, and their close contacts. Details of the protocol have been published previously^{44–46}. All participants provided written, informed consent and the protocol was approved by the Ethics Committee of the Maternidade Climério de Oliveira, Salvador, Brazil. All research was performed in accordance with relevant guidelines/regulations. Blood was collected in PAXgene tubes at diagnosis in TB cases (active TB, n = 51) and at enrolment in contacts (n = 99). Contacts with QFT IFN γ responses ≥ 0.35 IU/mL were considered latently *M.tb* infected (QFT-positive, n = 22) and uninfected if QFT IFN γ < 0.35 IU/mL (QFT-negative, n = 77).

Capillary blood cohort. Twenty adults (18 years or older) with recently diagnosed, microbiologically confirmed pulmonary TB, who were positive for either sputum MGIT or solid culture, Xpert MTB/RIF, Xpert MTB/RIF Ultra, or smear microscopy within the preceding two weeks and had received no more than two weeks of tuberculosis treatment were consecutively recruited from ongoing TB diagnostic and treatment studies at the South African Tuberculosis Vaccine Initiative (SATVI) field site. Twenty-nine healthy adults living in communities from the Cape Winelands region were also enrolled. Individuals with anaemia (haemoglobin less than

8.0 g/dl) or any other acute or chronic disease were excluded from both groups but no screening for HIV was performed. For each participant, 2.5 mL of venous blood was collected into PAXgene RNA tubes (Qiagen) while 20 μ L, 50 μ L or 100 μ L capillary blood was collected by fingerprick sequentially using 20 μ L or 50 μ L Minivettes (Sarstedt) without anti-coagulant and immediately transferred into 0.5 mL microtubes (Sarstedt) containing PAXgene fluid at an equivalent ratio to the manufacturer's recommendations, i.e. 1 μ L blood: 2.76 μ L PAXgene fluid. Samples were mixed by inversion (venous PAXgene tubes) or by flicking (capillary blood microtubes), incubated at room temperature for two hours, and stored at -40°C . Participants provided written informed consent and the protocol was approved by the Human Research Ethics committee of the Faculty of Health Sciences at the University of Cape Town (HREC 812/2017). All research was performed in accordance with relevant guidelines/regulations.

Catalysis treatment response cohort, "Catalysis" (TB treatment response in HIV-uninfected patients).

In total, 131 HIV-uninfected adults with newly diagnosed pulmonary TB, as confirmed by sputum culture, were recruited at primary healthcare clinics in Cape Town; 101 completed the study. Disease pathology was quantified by positron emission tomography and computerized tomography (PET-CT) imaging using ^{18}F FDG at baseline, week 4 and week 24. Total glycolytic activity index (TGAI) is a product of lesion volume and FDG uptake intensity and represents the total inflammatory burden, as previously provided^{18,19}. PAXgene tubes were collected prior to the start of treatment and at one, four, and 24 weeks after treatment initiation. Of the 101 sample sets sequenced for a transcriptomics analysis¹⁹, 84 patients met or exceeded the WHO definition for cure after the standard six-month treatment ("cures", had proven and then maintained sputum culture negativity by month 6). Amongst these, 70 had RNA available for qRT-PCR analysis. Eight patients did not achieve bacteriological cure (classified as "treatment failures", if the month 6 culture was still positive) and 7 had available RNA. None of the treatment failures achieved culture negativity at any time point during treatment and 7 had RNA for qRT-PCR analysis). The remaining 10 patients were probable cures (only final culture was negative) or unevaluable (treatment response ambiguous) and were not included in any analyses. Twenty-nine healthy controls were also enrolled from the same communities and 21 had RNA available for qRT-PCR analysis. All participants provided written, informed consent and the protocol was approved by the Stellenbosch University Human Research Ethics Committee (N10/01/013). All research was performed in accordance with relevant guidelines/regulations.

IMPRESS trial cohort (Recurrent TB treatment response in HIV-infected patients).

This study was an open-label, randomized controlled trial, "Improving Retreatment Success" (IMPRESS, clinicaltrials.gov, NTC02114684; SANCTR DOH-27-0414-4576), performed in Durban, KwaZulu-Natal²⁰. IMPRESS was designed to determine if a moxifloxacin-containing 24-week regimen, in which moxifloxacin was substituted for ethambutol, would improve TB retreatment outcomes relative to the standard TB treatment regimen. The trial enrolled adults with a previous history of TB disease who received a new diagnosis of drug-sensitive TB by positive Xpert MTB/RIF (Cepheid) or sputum smear microscopy, or both. Sputum samples were collected for culture testing every 2 weeks during the intensive phase of treatment and monthly thereafter until successful treatment completion. Whole blood was collected in PAXgene tubes at baseline, 7 days and 2, 6, 8 and 14 months after start of TB treatment. Sixty-three HIV-infected patients had RNA available and were included in the analyses (44 early converters, with sputum culture conversion before month 2; and 19 late converters, who converted after month 2). The IMPRESS protocol was reviewed and approved by the University of KwaZulu-Natal Biomedical Research Ethics Committee (BREC No. BFC029/13). The IMPRESS trial was also approved by the Medicines Control Council of South Africa (MCC Ref:20130510). All research was performed in accordance with relevant guidelines/regulations and all participants provided written, informed consent.

RNA extraction. RNA was manually extracted from collected PAXgene Blood RNA tubes (Qiagen) with the PAXgene blood RNA kit (Qiagen) according to the manufacturer's instructions or on an automated Tecan Freedom EVO 150 robotic platform with the Omega Maxwell SimplyRNA kit, using a modified protocol in a biosafety level 2 laboratory. Manually extracted RNA was stored at -80°C , and later used for transcriptomic analysis. For RNA extracted by robotic platform an aliquot was immediately used for cDNA synthesis. For the Peruvian cohort, RNA samples were extracted from 10^6 PBMCs using the RNeasy kit (Qiagen) according to manufacturers' instructions, and blinded, frozen aliquots of RNA were shipped to the University of Cape Town.

For the venous versus capillary blood comparison, RNA was isolated with the PAXgene blood RNA kit (Qiagen) according to the manufacturer's instructions with the following modifications: capillary blood samples were washed in 400 μ L water (instead of 4 ml) and homogenised by pipetting to avoid loss of the small pellet; venous and capillary samples were eluted in 80 μ L and 40 μ L of PAXgene blood RNA kit elution buffer, respectively.

Gene expression. cDNA synthesis and all PCR work was performed in a biosafety level 2 laboratory. cDNA was synthesized from extracted RNA using SuperScript II reverse transcriptase and pre-amplified using a pool of specific TaqMan primer-probe sets for microfluidic qRT-PCR. Gene expression of individual transcripts was then quantified by microfluidic qRT-PCR using either 96.96 or 192.24 Gene Expression chips on a BioMark HD (Fluidigm). An internal positive control sample was run on every chip to monitor inter-chip gene expression consistency.

Discovery of a parsimonious prognostic signature of TB disease risk. We sought to develop a PCR-based signature comprising a small ensemble of transcript pairs that each represent the ratio between one upregulated and one downregulated transcript in progressors, relative to controls, as described previously⁸. This pair-ratio ensemble format presents two advantages. Firstly, the up-down pairing provides a "self-standardisation"

function that eliminates the need for housekeeper transcript-based standardisation (normalisation) of RT-PCR cycle threshold values. Secondly, the ensemble of pairs provides robustness to the signature since a signature score can be calculated even if expression data for one transcript (and its pairs) is not available, due to a failed PCR reaction, for example.

Discovery of RISK6 signature of TB disease risk (Supplementary Figure 1a) was performed using all ACS cohort progressor/non-progressor samples collected within 360 days of TB disease diagnosis⁶. We first identified exon junctions that were differentially expressed in RNA-sequencing data from all progressors and matched non-progressors (published in⁶ and available on GEO: accession number GSE79362). We applied the random subsets approach, which randomly selects a partition of half the samples with a quarter of the features, to train support vector machines of all possible pairs of junctions using the Pair-Ratio approach. The Pair-Ratio approach pairs transcripts that are regulated in opposite directions in progressors and non-progressors. We identified transcript pairs that differentiated progressors and non-progressors with the highest sensitivity and specificity on the remaining partition of samples not used for fitting. This was repeated until pairs that comprised 84 unique exon junctions were identified, such that these could be conveniently assayed, along with 12 housekeeper (reference) transcripts, by microfluidic PCR in a 96-reaction format (Supplementary Figure 2).

Training RISK6, a prognostic PCR signature of TB disease risk. Taqman FAM-TAM primer-probe assays for each of the 84 exon junctions were used to measure expression of all transcripts by microfluidic qRT-PCR using samples from the entire ACS cohort. Delta Ct values were computed for each exon junction relative to the geometric mean of the 12 reference transcripts. To train the best parsimonious signature, we evaluated fit of pair-ratio ensembles consisting of either 10, 8, 6, 4 or 2 transcript pairs and evaluated their ability to differentiate between progressors and non-progressors. An appreciable drop in area under the receiver operating characteristic curve (ROC-AUC) was observed for the 2-transcript and 4-transcript ensembles, compared to the 6-, 8- and 10-transcript pair ratio ensembles (Supplementary Figure 2). As there was no significant difference in performance between the 6-, 8- and 10-transcript models, we selected the 6-primer model, which we termed RISK6, based on performance and smallest functional ensemble size.

Statistical analysis. All assays were performed on 96.96 or 192.24 Fluidigm gene expression chip format in duplicate. Each chip included an internal positive control and negative (no template) control. The following parameters were applied for extracting Ct values: Linear (Derivative) baseline correction, Quality Threshold of 0.65 and Auto (Global) for Ct Threshold Method using Fluidigm software version 3.1.3. Only replicate primer/probe assays with a $\geq 80\%$ pass rate across all samples within each gene expression chip, and samples with a $\geq 80\%$ replicate assay pass rate within each gene expression chip were retained for downstream analysis. To ensure high reproducibility between gene expression chips the RISK6 score obtained for the internal positive control sample run on each chip had to fall within a target range established on 11 historical gene expression chips, as 2 standard deviations of the mean RISK6 score. Intra-chip correlations between RISK6 assay Ct values of replicates for the internal positive control sample had to exceed >0.90 . RISK6 scores were computed for each replicate and the final score for a sample was considered as the average.

The RISK6 signature score is calculated as follows (R script available on Bitbucket: <https://bitbucket.org/satvi/risk6>):

1. Measure the cycle thresholds (Cts) for the 6 primer-probe assays listed in Supplementary Table 1, by qRT-PCR.
2. For each of the 9 transcript pairs, compute the difference in raw Ct, which produces the log-transformed ratio of expression.
3. Compare the measured ratio to ratios in a look-up table for the given pair of transcripts.
4. Assign a corresponding score in the look-up table to the ratio. If the measured ratio is larger than all ratios in the relevant column of the look-up table, then assign a score of 1 to the ratio.
5. Compute the average over the scores generated from the set of pairs. If any assays failed on the sample, compute the average score over all ratios not including the failed assays. The resulting average is the final score for that sample.

There is considerable interest in the biosignature field to apply such signatures to publicly available microarray or RNA-sequencing data^{11,47}. Microarray datasets were obtained programmatically from Gene Expression Omnibus and processed using the MetaIntegrator R package, as previously described⁴⁷. RISK6geo scores were computed from the gene-level summarized intensities. RISK6 scores can be computed from \log_2 -transformed microarray or RNA-sequencing data using the formula:

$$\text{RISK6geo score} = \text{geometric mean}(\text{GBP2}, \text{FCGR1B}, \text{SERPING1}) \\ - \text{geometric mean}(\text{TUBGCP6}, \text{TRMT2A}, \text{SDR39U1})$$

where normalized \log_2 -transformed mean fluorescence intensity or normalized read count values of GBP2, FCGR1B, SERPING1, TUBGCP6, TRMT2A and SDR39U1 are used.

RISK6 scores can also be computed using this method from qRT-PCR data using the formula:

$$\text{RISK6geo score} = \text{geometric mean}(\text{TUBGCP6}, \text{TRMT2A}, \text{SDR39U1}) \\ - \text{geometric mean}(\text{GBP2}, \text{FCGR1B}, \text{SERPING1})$$

where raw Ct values of GBP2, FCGR1B, SERPING1, TUBGCP6, TRMT2A and SDR39U1 are used. Comparative performance characteristics of the RISK6 and RISK6geo signatures for the different cohorts in this study are shown in Supplementary Table 2.

qRT-PCR gene expression data was quality controlled using scripts generated in R and signature scores were calculated. All RISK6 scores, with the exception of those in the discovery cohort, were generated by blinded laboratory personnel. Only once RISK6 score results were locked down and, where appropriate, shared among collaborators, were group allocations unblinded for performance analyses. ROC AUCs were generated and compared using the pROC⁴⁸ and verification⁴⁹ packages in R. Statistical analyses were done using Mann Whitney U for differences between two groups, Wilcoxon ranked sum and Kruskal-Wallis tests for differences between three groups in GraphPad Prism v8. To generate spline plots that show temporal changes in transcript expression between adolescent progressors and non-progressors, we computed log₂ fold change values between progressor and non-progressors transcript abundance (measured by RNA-sequencing) as previously described⁴¹, and modeled these as a nonlinear function of TimeToDiagnosis for the entire adolescent progressor/non-progressor cohort using the smooth.spline function in R with three degrees of freedom. Ninety-nine percent confidence intervals for the temporal trends were computed by performing 2000 iterations of spline fitting after bootstrap resampling from the full dataset. The median difference and 95% CIs in expression of RISK6 signature genes was computed from 1000 bootstrapped median Ct values between HIV+ and HIV- individuals. Genes with 95% CI bounds above zero were considered significant.

Data availability

The RISK6 scores and associated clinical data for all cohorts are in Supplementary Tables 5–13.

Received: 8 September 2019; Accepted: 27 April 2020;

Published online: 25 May 2020

References

- World Health Organization. *The end TB strategy*. (2014).
- Churchyard, G. J. A stratified approach to tuberculosis treatment. *Nature medicine* **24**, 1639–1641 (2018).
- Imperial, M. Z. *et al.* A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nature medicine* **24**, 1708–1715 (2018).
- World Health Organization. FIND. *Development of a Target Product Profile (TPP) and a framework for evaluation for a test for predicting progression from tuberculosis infection to active disease*. (2017).
- World Health Organization. *High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting*. (2014).
- Zak, D. E. *et al.* A blood RNA signature for tuberculosis disease risk: a prospective cohort study. *The Lancet* **387**, 2312–2322 (2016).
- Darboe, F. *et al.* Diagnostic performance of an optimized transcriptomic signature of risk of tuberculosis in cryopreserved peripheral blood mononuclear cells. *Tuberculosis (Edinburgh, Scotland)* **108**, 124–126 (2018).
- Suliman, S. *et al.* Four-gene Pan-African Blood Signature Predicts Progression to Tuberculosis. *American journal of respiratory and critical care medicine* **197**, 1198–1208 (2018).
- Sweeney, T. E., Braviak, L., Tato, C. M. & Khatri, P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. *Lancet Respir Med* **4**, 213–224 (2016).
- Maertzdorf, J. *et al.* Concise gene signature for point-of-care classification of tuberculosis. *EMBO molecular medicine* **8**, 86–95 (2016).
- Warsinske, H., Vashisht, R. & Khatri, P. Host-response-based gene signatures for tuberculosis diagnosis: A systematic comparison of 16 signatures. *PLoS Med.* **16**, e1002786 (2019).
- MacLean, E. *et al.* A systematic review of biomarkers to detect active tuberculosis. *Nat Microbiol* **4**, 748–758 (2019).
- World Health Organization. *Global Tuberculosis Report 2019*. (2019).
- Manngo, P. M. *et al.* Prospective evaluation of host biomarkers other than interferon gamma in QuantiFERON Plus supernatants as candidates for the diagnosis of tuberculosis in symptomatic individuals. *The Journal of infection* **79**, 228–235 (2019).
- Chegou, N. N. *et al.* Diagnostic performance of a seven-marker serum protein biosignature for the diagnosis of active TB disease in African primary healthcare clinic attendees with signs and symptoms suggestive of TB. *Thorax* **71**, 785–794 (2016).
- Chegou, N. N. *et al.* Africa-wide evaluation of host biomarkers in QuantiFERON supernatants for the diagnosis of pulmonary tuberculosis. *Scientific reports* **8**, 2675–12 (2018).
- Malherbe, S. T. *et al.* A semi-automatic technique to quantify complex tuberculous lung lesions on 18F-fluorodeoxyglucose positron emission tomography/computerised tomography images. *EJNMMI Res* **8**, 55–14 (2018).
- Malherbe, S. T. *et al.* Persisting positron emission tomography lesion activity and Mycobacterium tuberculosis mRNA after tuberculosis cure. *Nature medicine* **22**, 1094–1100 (2016).
- Thompson, E. G. *et al.* Host blood RNA signatures predict the outcome of tuberculosis treatment. *Tuberculosis (Edinburgh, Scotland)* **107**, 48–58 (2017).
- Perumal, R. *et al.* A moxifloxacin-based regimen for the treatment of recurrent drug-sensitive pulmonary tuberculosis: An open-label randomised controlled trial. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **362**(5), 13 (2019).
- Deeks, S. G., Tracy, R. & Douek, D. C. Systemic effects of inflammation on health during chronic HIV infection. *Immunity* **39**, 633–645 (2013).
- Fraser, C. *et al.* Virulence and pathogenesis of HIV-1 infection: an evolutionary perspective. *Science* **343**, 1243727–1243727 (2014).
- Joshi, A. *et al.* HIV-1 Env Glycoprotein Phenotype along with Immune Activation Determines CD4 T Cell Loss in HIV Patients. *J. Immunol.* **196**, 1768–1779 (2016).
- McNab, F., Mayer-Barber, K., Sher, A., Wack, A. & O'Garra, A. Type I interferons in infectious disease. *Nat. Rev. Immunol.* **15**, 87–103 (2015).
- Lehmann, C. *et al.* Increased interferon alpha expression in circulating plasmacytoid dendritic cells of HIV-1-infected patients. *J. Acquir. Immune Defic. Syndr.* **48**, 522–530 (2008).
- Klatt, N. R., Chomont, N., Douek, D. C. & Deeks, S. G. Immune activation and HIV persistence: implications for curative approaches to HIV infection. *Immunol. Rev.* **254**, 326–342 (2013).
- Andrews, J. R. *et al.* The dynamics of QuantiFERON-TB gold in-tube conversion and reversion in a cohort of South African adolescents. *American journal of respiratory and critical care medicine* **191**, 584–591 (2015).
- Gupta, R. K. *et al.* Concise whole blood transcriptional signatures for incipient tuberculosis: A systematic review and patient-level pooled meta-analysis. *bioRxiv* 668137, <https://doi.org/10.1101/668137> (2019).

29. Sambarey, A. *et al.* Meta-analysis of host response networks identifies a common core in tuberculosis. *NPJ Syst Biol Appl* **3**, 4 (2017).
30. Onozaki, I. *et al.* National tuberculosis prevalence surveys in Asia, 1990–2012: an overview of results and lessons learned. *Trop. Med. Int. Health* **20**, 1128–1145 (2015).
31. Calligaro, G. L. *et al.* Effect of new tuberculosis diagnostic technologies on community-based intensified case finding: a multicentre randomised controlled trial. *Lancet Infect Dis* **17**, 441–450 (2017).
32. Dowdy, D. W. *et al.* Designing and Evaluating Interventions to Halt the Transmission of Tuberculosis. *J. Infect. Dis* **216**, S654–S661 (2017).
33. Warsinske, H. C. *et al.* Assessment of Validity of a Blood-Based 3-Gene Signature Score for Progression and Diagnosis of Tuberculosis, Disease Severity, and Treatment Response. *JAMA Netw Open* **1**, e183779 (2018).
34. Singhania, A. *et al.* A modular transcriptional signature identifies phenotypic heterogeneity of human tuberculosis infection. *Nat Commun* **9**, 2308 (2018).
35. Herbeval, J.-P. & Shearer, G. M. HIV-1 immunopathogenesis: how good interferon turns bad. *Clinical immunology (Orlando, Fla)* **123**, 121–128 (2007).
36. Esmail, H. *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* **115**, E964–E973 (2018).
37. Cliff, J. M. *et al.* Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. *J. Infect. Dis* **207**, 18–29 (2013).
38. Bloom, C. I. *et al.* Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. *PLoS ONE* **7**, e46191 (2012).
39. SEQC/MAQC-III Consortium. A comprehensive assessment of RNA-seq accuracy, reproducibility and information content by the Sequencing Quality Control Consortium. *Nat. Biotechnol.* **32**, 903–914 (2014).
40. Leek, J. T. *et al.* Tackling the widespread and critical impact of batch effects in high-throughput data. *Nat. Rev. Genet.* **11**, 733–739 (2010).
41. Scriba, T. J. *et al.* Sequential inflammatory processes define human progression from M. tuberculosis infection to tuberculosis disease. *PLoS pathogens* **13**, e1006687 (2017).
42. Naidoo, A. *et al.* Effect of rifampicin and efavirenz on moxifloxacin concentrations when co-administered in patients with drug-susceptible TB. *J. Antimicrob. Chemother* **72**, 1441–1449 (2017).
43. Duffy, F. J. *et al.* A Serum Circulating miRNA Signature for Short-Term Risk of Progression to Active Tuberculosis Among Household Contacts. *Front Immunol* **9** (2018).
44. Hamilton, C. D. *et al.* RePORT International: Advancing Tuberculosis Biomarker Research Through Global Collaboration. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America* **61Suppl 3**, S155–9 (2015).
45. Geadas, C. *et al.* Advances in basic and translational tuberculosis research: Proceedings of the first meeting of RePORT international. in **102**, 55–67 (2017).
46. van der Heijden, Y. F. *et al.* Building capacity for advances in tuberculosis research; proceedings of the third RePORT international meeting. *Tuberculosis (Edinburgh, Scotland)* **113**, 153–162 (2018).
47. Haynes, W. A. *et al.* Empowering multi-cohort gene expression analysis to increase reproducibility. *Pac Symp Biocomput* **22**, 144–153 (2017).
48. Robin, X. *et al.* Display and Analyze ROC Curves [R package pROC version 1.13.0]. (2013).
49. Pocerlich, M. Package 'verification'. In CRAN.R-project.org Internet (2015).

Acknowledgements

Support for this study was provided by the Bill and Melinda Gates Foundation (grants OPP1023483, OPP1065330, and Grand Challenges in Global Health (GC6-74 grant 37772)) and the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council with funds received from the South African Department of Science and Technology. The ACS study was also supported by BMGF GC12 (grant 37885) for QuantiFERON testing. The IMPRESS trial was supported by the European and Developing Countries Clinical Trials Partnership (EDCTP: TA.2011.40200.044), and by Bayer Healthcare (moxifloxacin donation). The AE-TBC and ScreenTB projects were funded by the EDCTP with grant numbers IP_2009_32040 and DRIA2014-311 respectively. RePORT Brazil is supported by the Departamento de Ciência e Tecnologia (DECIT) - Secretaria de Ciência e Tecnologia (SCTIE) - Ministério da Saúde (MS), Brazil [25029.000507/2013-07] the National Institutes of Allergy and Infectious Diseases [U01-AI069923], and CRDF Global [DAA3-18-64151, DAA3-18-64152 and DAA3-18-64153]. Work in Peru was supported by the Tuberculosis Research Unit Network (U19 AI111224). SCM received training in research that was supported by the Fogarty International Center of the National Institutes of Health under Award Number D43 TW010559. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. F.D. was supported by the Margaret McNamara educational grant for women in developing countries. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author contributions

A.P.N., E.T., D.Z. and T.J.S. conceived the study. A.P.N., S.K.M., W.A.H., S.H.E.K., J.W., R.W., I.V.R., B.M., M.Murray, B.B.A., T.R.S., J.S., G.W., K.N., N.P., M.H., D.Z. and T.J.S. implemented clinical studies, raised funds and/or provided the resources. A.P.N., E.T., S.K.M., S.C.M., S.S., N.N.C., S.T.M., F.D., M.E., N.B., M.Murphy, I.V.R. and B.B.A. processed samples, performed the experiments, and analyzed the data. A.P.N., S.S., N.N.C., N.B., M.F., M.M., C.M. provided operational and/or project management. A.P.N., E.T., S.K.M., S.C.M., S.S., N.N.C., S.T.M., F.D., J.W., I.V.R., B.M., M.Murray, B.B.A., T.R.S., G.W., M.H., D.Z. and T.J.S. interpreted the results. A.P.N., E.T., S.K.M., S.C.M., S.S., D.Z. and T.J.S. wrote the manuscript. Members of the Adolescent Cohort Study team, GC6-74 Consortium, the SATVI Clinical and Laboratory Team, The ScreenTB and AE-TBC teams, CAPRISA IMPRESS team, RePORT Brazil Consortium and Peruvian Household Contacts Cohort study group performed clinical studies and processed samples. All authors have read and approved the manuscript.

Competing interests

A.P.N., E.T., W.A.H., D.Z. and T.J.S. are co-inventors of a patent on RISK6. All other authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-65043-8>.

Correspondence and requests for materials should be addressed to T.J.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020

The Adolescent Cohort Study team

Fazlin Kafaar¹, Leslie Workman¹, Humphrey Mulenga¹, E. Jane Hughes¹, Onke Xasa¹, Ashley Veldsman¹, Yolundi Cloete¹, Deborah Abrahams¹, Sizulu Moyo¹, Sebastian Gelderbloem¹, Michele Tameris¹, Hennie Geldenhuys¹, Rodney Ehrlich¹⁵, Suzanne Verver¹⁶ & Larry Geiter¹⁷

¹⁵School of Public Health and Family Medicine, University of Cape Town, Cape Town, South Africa. ¹⁶KNCV Tuberculosis Foundation, The Hague, and Amsterdam Institute of Global Health and Development, Academic Medical Centre, Amsterdam, The Netherlands. ¹⁷Aeras, Rockville, MD, USA.

The GC6-74 Consortium

Gillian F. Black⁴, Gian van der Spuy⁴, Kim Stanley⁴, Magdalena Kriel⁴, Nelita Du Plessis⁴, Nonhlanhla Nene⁴, Teri Roberts⁴, Leanie Kleynhans⁴, Andrea Gutschmidt⁴, Bronwyn Smith⁴, Andre G. Loxton⁴, Gerhardus Tromp⁴, David Tabb⁴, Tom H. M. Ottenhoff¹⁸, Michel R. Klein¹⁸, Marielle C. Haks¹⁸, Kees L. M. C. Franken¹⁸, Annemieke Geluk¹⁸, Krista E. van Meijgaarden¹⁸, Simone A Joosten¹⁸, W. Henry Boom¹⁹, Bonnie Thiel¹⁹, Harriet Mayanja-Kizza²⁰, Moses Joloba²⁰, Sarah Zalwango²⁰, Mary Nsereko²⁰, Brenda Okwera²⁰, Hussein Kisingo²⁰, Shreemanta K. Parida⁵, Robert Golinski⁵, Jeroen Maertzdorf⁵, January Weiner 3rd⁵, Marc Jacobson⁵, Hazel Dockrell²¹, Steven Smith²¹, Patricia Gorak-Stolinska²¹, Yun-Gyoung Hur²¹, Maeve Lalor²¹, Ji-Sook Lee²¹, Amelia C. Crampin²², Neil French²², Bagrey Ngwira²², Anne Ben-Smith²², Kate Watkins²², Lyn Ambrose²², Felanji Simukonda²², Hazzie Mvula²², Femia Chilongo²², Jacky Saul²², Keith Branson²², Hassan Mahomed¹, E. Jane Hughes¹, Onke Xasa¹, Ashley Veldsman¹, Katrina Downing¹, Humphrey Mulenga¹, Brian Abel¹, Mark Bowmaker¹, Benjamin Kagina¹, William Kwong Chung¹, Jerry Sadoff¹⁷, Donata Sizemore¹⁷, S Ramachandran¹⁷, Lew Barker¹⁷, Michael Brennan¹⁷, Frank Weichold¹⁷, Stefanie Muller¹⁷, Larry Geiter¹⁷, Desta Kassa²³, Almaz Abebe²³, Tsehayenesh Mesele²³, Belete Tegbaru²³, Debbie van Baarle²⁴, Frank Miedema²⁴, Rawleigh Howe²⁵, Adane Mihret²⁵, Abraham Aseffa²⁵, Yonas Bekele²⁵, Rachel Iwnetu²⁵, Mesfin Tafesse²⁵, Lawrence Yamuah²⁵, Martin Ota¹², Philip Hill¹², Richard Adegbola¹², Tuman Corrah¹², Martin Antonio¹², Toyin Togun¹², Ifedayo Adetifa¹², Simon Donkor¹², Peter Andersen²⁶, Ida Rosenkrands²⁶, Mark Doherty²⁶, Karin Weldingh²⁶, Gary Schoolnik²⁷, Gregory Dolganov²⁷ & Tran Van²⁷

¹⁸Department of Infectious Diseases, Leiden University Medical Centre, Leiden, The Netherlands. ¹⁹Tuberculosis Research Unit, Department of Medicine, Case Western Reserve University School of Medicine and University Hospitals Case Medical Center, Cleveland, Ohio, USA. ²⁰Department of Medicine and Department of Microbiology, College of Health Sciences, Faculty of Medicine, Makerere University, Kampala, Uganda. ²¹Department of Immunology and Infection, Faculty of Infectious and Tropical Diseases, London School of Hygiene & Tropical Medicine, London, United Kingdom. ²²Karonga Prevention Study, Chilumba, Malawi. ²³Ethiopian Health & Nutrition Research Institute, Addis Ababa, Ethiopia. ²⁴University Medical Centre, Utrecht, The Netherlands. ²⁵Armauer Hansen Research Institute, Addis Ababa, Ethiopia. ²⁶Department of Infectious Disease Immunology, Statens Serum Institute, Copenhagen, Denmark. ²⁷Department of Microbiology and Immunology, Stanford University, Stanford, California, USA.

The SATVI Clinical and Laboratory Team

Denis Arendsen¹, Hadn Africa¹, Veronica Baartman¹, Elizabeth Filander¹, Cynthia Gwintsa¹, Simbarashe Mabwe¹, Lebohang Makhethhe¹, Miriam Moses¹, Raida Onrust¹, Michele van Rooyen¹, Marcia Steyn¹ & Habibullah Valley¹

The ScreenTB Consortium

Petri Ahlers⁴, Gian van der Spuy⁴, Ilana van Rensburg⁴, Hygon Mutavhatsindi⁴, Portia Manngo⁴, Kim Stanley⁴, Andriette Hiemstra⁴, Shirley McAnda⁴, Joseph Mendy¹², Awa Gindeh¹², Georgetta Mbayo¹², Ebrima Trawally¹², Olumuyiwa Owolabi¹², Harriet Mayanja-Kizza²⁰, Mary Nsereko²⁰, Anna-Rita Namuganga²⁰, Saudah Nambiru Kizito²⁰, Adane Mihret²⁵, Sosina Ayalew²⁵, Rawleigh Howe²⁵, Azab Tarekegne²⁵, Bamlak Tessema²⁵, Emmanuel Nepolo²⁸, Joseph Sheehama²⁸, Gunar Gunther²⁸, Azaria Diergaardt²⁸, Uapa Pazvakavambwa²⁸, Hazel Dockrell²¹, Tom Ottenhoff¹⁸, Elisa Tjon Kon Fat¹⁸, Shannon Herdigein¹⁸, Paul Corstjens¹⁸ & Annemieke Geluk¹⁸

²⁸University of Namibia, Windhoek, Namibia.

The AE-TBC Consortium

Magdalena Kriel⁴, Gian van der Spuy⁴, Andre G. Loxton⁴, Kim Stanley⁴, Belinda Kriel⁴, Leigh A Kotzé⁴, Dolapo O. Awoniyi⁴, Elizna Maasdorp⁴, Olumuyiwa Owolabi¹², Abdou Sillah¹², Joseph Mendy¹², Awa Gindeh¹², Simon Donkor¹², Toyin Togun¹², Martin Ota¹², Harriet Mayanja-Kizza²⁰, Ann Ritah Namuganga²⁰, Grace Muzanye²⁰, Mary Nsereko²⁰, Pierre Peters²⁰, Marieta van der Vyver²⁸, Faustina N. Amutenya²⁸, Josefina N. Nelongo²⁸, Lidia Monye²⁸, Jacob A. Sheehama²⁸, Scholastica Ipinge²⁸, Amelia C. Crampin²², Felanji Simukonda²², Alemayehu Amberbir²², Femia Chilongo²², Rein Houben²², Desta Kassa²³, Atsbeha Gebrezgabher²³, Getnet Mesfin²³, Yohannes Belay²³, Gebremedhin Gebremichael²³, Yodit Alemayehu²³, Rawleigh Howe²⁵, Adane Mihret²⁵, Yonas Bekele²⁵, Bamlak Tessema²⁵, Lawrence Yamuah²⁵, Tom H. M. Ottenhoff¹⁸, Annemieke Geluk¹⁸, Kees L. M. C. Franken¹⁸, Paul L. A. M. Corstjens¹⁸, Elisa M. Tjon Kon Fat¹⁸, Claudia J. de Dood¹⁸, Jolien J. van der Ploeg-van Schip¹⁸, Ida Rosenkrands²⁶, Claus Aagaard²⁶, Maria M. Esterhuysen⁵, Jacqueline M. Cliff²¹ & Hazel M. Dockrell²¹

The RePORT Brazil Team

Juan M. Cubillos-Angulo¹⁰, Kiyoshi F. Fukutani¹⁰, Laise Paixão¹⁰, Ricardo Khouri¹⁰, Sayonara Melo¹⁰, Alice Andrade¹⁰, Jéssica Rebouças-Silva¹⁰, Hayna Malta¹⁰, Artur T. L. Queiroz¹⁰, Valeria C. Rolla²⁹, Solange Cavalcante²⁹, Betina Durovni³⁰, Marcelo Cordeiro-Santos³¹, Afranio Kritski³², José R. Lapa e Silva³² & Marina C. Figueiredo¹¹

²⁹Instituto Nacional de Infectologia Evandro Chagas, Fundação Oswaldo Cruz, Rio de Janeiro, Brazil. ³⁰Secretaria Municipal de Saúde do Rio de Janeiro, Coordenação de Doenças Transmissíveis, Rio de Janeiro, Brazil. ³¹Fundação de Medicina Tropical Doutor Heitor Vieira Dourado, Manaus, Brazil. ³²Hospital Universitário Clementino Fraga Filho, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

Peruvian Household Contacts Cohort Team

Kattya Lopez Tamara³, Kattya Lopez Tamara³³, Segundo R. León³³ & Leonid Lecca Garcia³³

³³Socios En Salud, Lima, Peru.

The CAPRISA IMPRESS team

Dhineshree Govender^{13,14}, Razia Hassan-Moosa^{13,14}, Anushka Naidoo^{13,14}, Rochelle Adams^{13,14}, Natasha Samsunder^{13,14} & Lara Lewis^{13,14}