#### **1** Prospective validation of host transcriptomic biomarkers for pulmonary

#### 2 tuberculosis by real-time PCR

3

Simon C. Mendelsohn, MBChB<sup>1</sup>, Stanley Kimbung Mbandi, PhD<sup>1</sup>, Andrew Fiore-Gartland, PhD<sup>2</sup>,
Adam Penn-Nicholson, PhD<sup>1</sup>, Munyaradzi Musvosvi, PhD<sup>1</sup>, Humphrey Mulenga, MPH<sup>1</sup>, Michelle
Fisher, PhD<sup>1</sup>, Katie Hadley, PhD<sup>1</sup>, Mzwandile Erasmus, BSc<sup>1</sup>, Onke Nombida, GT<sup>1</sup>, Michèle
Tameris, MBChB<sup>1</sup>, Gerhard Walzl, MBChB<sup>3</sup>, Kogieleum Naidoo, PhD<sup>4,5</sup>, Gavin Churchyard,
PhD<sup>6,7,8</sup>, Mark Hatherill, MD<sup>1</sup>, Thomas J. Scriba, DPhil<sup>1</sup>, The Cross-sectional TB Cohort Study
Team\*, and The CORTIS Study Team\*

- 10
- <sup>11</sup> <sup>1</sup>South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and Molecular
- 12 Medicine, and Division of Immunology, Department of Pathology, University of Cape Town, 7925,
- 13 Cape Town, South Africa.

<sup>14</sup> <sup>2</sup>Vaccine and Infectious Disease Division, Fred Hutchinson Cancer Research Center, Seattle, WA

- 15 98109, USA.
- <sup>16</sup> <sup>3</sup>DST/NRF Centre of Excellence for Biomedical TB Research; South African Medical Research
- 17 Council Centre for TB Research; Division of Molecular Biology and Human Genetics, Department
- 18 of Biomedical Sciences, Faculty of Medicine and Health Sciences, Stellenbosch University, 7505,
- 19 Cape Town, South Africa.
- <sup>4</sup>Centre for the AIDS Programme of Research in South Africa (CAPRISA), 4001, Durban, South
   Africa.
- <sup>22</sup> <sup>5</sup>MRC-CAPRISA HIV-TB Pathogenesis and Treatment Research Unit, Doris Duke Medical
- 23 Research Institute, University of KwaZulu-Natal, 4001, Durban, South Africa.
- <sup>6</sup>The Aurum Institute, 2194, Johannesburg, South Africa.
- <sup>25</sup> <sup>7</sup>School of Public Health, University of Witwatersrand, 2193, Johannesburg, South Africa.
- <sup>26</sup> <sup>8</sup>Department of Medicine, Vanderbilt University, Nashville, TN 37232, USA.
- 27

- 29
- 30 **Word count:** 6,821 words (48,585 characters including spaces)
- 31 Running title: Blood mRNA biomarkers for pulmonary TB
- 32
- 33 Corresponding author: Thomas J Scriba
- 34 Email: Thomas.Scriba@uct.ac.za Telephone: +27 (0) 21 406 6427
- 35 **Postal address:** South African Tuberculosis Vaccine Initiative, Institute of Infectious Disease and
- 36 Molecular Medicine, Division of Immunology, Department of Pathology, University of Cape Town,
- 37 Anzio Road, Observatory, 7925, Cape Town, South Africa

<sup>28 \*</sup> See supplement

#### 38 Abstract

39 We tested performance of host-blood transcriptomic tuberculosis (TB) signatures for active case-40 finding. Among 20,207 HIV-uninfected and 963 HIV-infected adults screened; 2,923 and 861 were 41 enrolled from five South African communities. Eight signatures were measured by microfluidic RT-42 qPCR and participants were microbiologically-investigated for pulmonary TB at baseline, and 43 actively surveilled for incident disease through 15 months. Diagnostic AUCs for 61 HIV-uninfected 44 (weighted-prevalence 1.1%) and 10 HIV-infected (prevalence 1.2%) prevalent TB cases for the 8 45 signatures were 0.63–0.79 and 0.65–0.88, respectively. Thereafter, 24 HIV-uninfected and 9 HIV-46 infected participants progressed to incident TB (1.1 and 1.0 per 100 person-years, respectively). 47 Prognostic AUCs through 15-months follow-up were 0.49-0.66 and 0.54-0.81, respectively. 48 Prognostic performance for incident TB occurring within 6-12 months in HIV-negative participants 49 was higher for all signatures. None of the signatures met WHO Target Product Profile criteria for a 50 triage test to diagnose asymptomatic TB; most signatures met the criteria for symptomatic TB. 51 Prognostic accuracy of most signatures for incident TB within six months of testing met the criteria 52 for an incipient TB test. 53

- 54 Keywords
- 55 Diagnostic / prognostic/ transcriptional / signature / Tuberculosis

#### 56 Introduction

57 An estimated 2.9 million people with tuberculosis (TB) disease went undiagnosed or unreported in 58 2019<sup>1</sup>, partly because passive case-finding strategies rely primarily on self-presentation of people 59 for healthcare when they experience TB symptoms, and dependence on sputum for 60 microbiological testing. TB prevalence surveys in Africa and Asia demonstrate that approximately 61 50% of TB cases are asymptomatic, highlighting a major blind-spot in the symptom-dependent 62 case-finding strategies, which may perpetuate significant Mycobacterium tuberculosis (Mtb) transmission.<sup>2-5</sup> Currently available triage tests and passive case-finding strategies for symptomatic 63 64 disease are thus inadequate to find these missing cases among people who do seek healthcare. 65 Effective active case-finding tools for screening of asymptomatic individuals in the community are 66 urgently needed for earlier detection, further microbiological testing, and treatment, in order to 67 interrupt the spread of *Mtb*.

68

69 Computer-aided detection for automated interpretation of digital chest radiographs<sup>6</sup> is one such 70 approach for rapid and cheap mass screening, particularly in resource-constrained settings. Mobile 71 sputum Xpert MTB/RIF screening also performs well as a community-based case-finding strategy, 72 but requires significant resources and is dependent on availability of a sputum sample.<sup>7</sup> Tests for 73 detection of *Mtb*-exposed individuals at high risk of progression to TB disease, particularly in high 74 burden settings, are also urgently needed to guide TB preventive therapy (TPT). Interferon-y 75 release assays (IGRA) and tuberculin skin testing lack specificity in endemic settings due to high 76 rates of prior Mtb exposure, which result in both overtreatment of those unlikely to progress to 77 disease and missed "incipient" cases.<sup>8</sup> IGRA are also laborious, and require laboratory facilities 78 and skilled technicians.

79

80 We previously developed a predictive correlate of risk signature, RISK11, for diagnosis of TB disease and for identification of individuals at high risk of progression to active TB disease.<sup>9,10</sup> 81 Diagnostic and prognostic performance of RISK11 as a PCR-based screening test was 82 83 prospectively validated in HIV-uninfected adults<sup>11</sup> and people living with HIV (PLHIV)<sup>12</sup>. This 84 signature is not ideal for further development due to its large size (48 primer-probe assays). 85 Fortunately, several refined, parsimonious diagnostic and prognostic transcriptomic signatures, which are more amenable to translation to point-of-care testing, have been developed.<sup>13-16</sup> 86 87 However, most transcriptomic signature discovery and validation studies use a case-control design 88 with stringent exclusion criteria, resulting in spectrum bias and best-case discriminative 89 performance, which is not reproducible in prospective studies in heterogenous real-world settings. 90 Transcriptomic signatures are predominantly measured by microarray or RNA-sequencing, 91 techniques which are not suitable for use at the point of care. There are also few studies which have performed unbiased head-to-head comparisons of signature performance<sup>13-15</sup> and none for 92

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

95

96 We recently showed that HIV and upper respiratory tract viruses result in an increase in RISK11 97 score and hypothesised that this was due to the upregulation of interferon-stimulated genes (ISG) as part of an inflammatory response typically induced in TB and viral infections.<sup>10,12,17</sup> Further, 98 99 RISK11 was able to differentiate individuals with TB disease from those with bacterial or no upper 100 respiratory tract microorganisms, but was not able to discriminate between those with TB from 101 those with upper respiratory tract viruses. Most TB transcriptomic signatures primarily detect 102 elevated ISG expression. The effect of HIV and respiratory viruses on scores and performance of 103 these other signatures, including those which do not encompass ISGs such as the Thompson5 (RESPONSE5) signature<sup>18</sup>, has not been adequately explored. 104

105

106 In this paper we report the results of a multi-centre observational study, which prospectively tested 107 the performance of eight parsimonious host-blood TB transcriptomic signatures measured by 108 microfluidic real-time (RT) qPCR, for diagnosing prevalent pulmonary TB and predicting 109 progression to incident TB disease, in a predominantly asymptomatic community cohort of HIV-110 uninfected and HIV-infected South African adults. As secondary aims, we compared signature 111 diagnostic performance in symptomatic and asymptomatic participants, and benchmarked 112 parsimonious signature accuracy against RISK11 and the World Health Organization (WHO) Target Product Profile (TPP) criteria for triage and prognostic TB tests.<sup>19,20</sup> As an exploratory aim, 113 114 we also explored the effect of HIV and upper respiratory tract microorganisms on transcriptomic 115 signature scores.

#### 116 Methods

#### 117

#### 118 Study design and participants

119 In this diagnostic and prognostic accuracy study, we evaluated parsimonious blood-based host 120 transcriptomic signature performance in four South African cohorts (Figure S1). We first translated 121 the signatures to a RT-qPCR platform and validated performance in a case-control study of HIV-122 uninfected and HIV-infected individuals with and without pulmonary TB, the Cross-sectional TB 123 Cohort (CTBC) study (Figure S1A). We then prospectively tested signature performance in HIV-124 uninfected volunteers enrolled in the CORTIS-01 trial (ClinicalTrials.gov: NCT02735590; Figure 125 **S1B**) and PLHIV enrolled in the observational CORTIS-HR study (Figure S1D). Finally, we 126 evaluated the effects of upper respiratory microorganisms on transcriptomic signatures in a subset 127 of HIV-uninfected participants screened for the CORTIS-01 trial (Figure S1C). This study is 128 reported in accordance with the Standards for Reporting Diagnostic Accuracy Studies initiative 129 guidelines<sup>21</sup> and the study protocol was approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics Committee (HREC 812/2017). Written informed consent 130 131 was obtained from all participants and all experiments conformed to the principles set out in the 132 WMA Declaration of Helsinki and the US Department of Health and Human Services Belmont 133 Report.

134

#### 135 Cross-sectional TB Cohort

The CTBC case-control study was previously described.<sup>22</sup> Briefly, adults aged ≥18 years with 136 137 newly diagnosed pulmonary TB confirmed by sputum Xpert MTB/RIF (Cepheid, Sunnyvale, CA, 138 USA) and/or liquid mycobacterial culture (Mycobacteria Growth Indicator Tube [MGIT], BACTEC, 139 Beckton Dickinson, Franklin Lakes, NJ, USA) were enrolled at primary healthcare clinics in 140 Worcester and Masiphumelele, South Africa. Healthy, asymptomatic community controls were 141 recruited from the same areas. HIV-infection was diagnosed with the Determine HIV1/2 test (Alere, 142 Waltham, MA, USA). All participants provided written, informed consent and the protocols were 143 approved by the University of Cape Town Faculty of Health Sciences Human Research Ethics 144 Committee (HREC 126/2006 and 288/2008).

145

#### 146 Correlate of Risk Targeted Intervention Study (CORTIS)

The CORTIS-01 clinical trial in HIV-uninfected individuals and CORTIS-HR observational study in PLHIV have previously been described.<sup>11,12</sup> Briefly, healthy adult volunteers without clinical suspicion of TB, residing in five TB endemic communities in South Africa (Durban, Klerksdorp, Ravensmead, Rustenburg, and Worcester), were recruited through word-of-mouth, house-tohouse visits, and liaison with non-governmental organisations. Recruitment did not target symptomatic individuals seeking health care or other high-risk groups. Eligible participants aged 18–59 years were without comorbidities (except for HIV) and did not have known TB disease, or 154 household exposure to individuals with multi-drug resistant TB, within the prior three years. The 155 RISK11 signature was measured at enrolment, and a pre-specified score threshold of 60% was 156 used to differentiate RISK11-positive from RISK11-negative participants. The HIV-uninfected 157 CORTIS-01 population was enriched for participants at high risk of TB disease through enrolment 158 of all eligible RISK11-positive individuals, and randomisation of RISK11-negative participants to 159 enrolment or non-enrolment at a 1:7-9 ratio. RISK11-positive participants were randomised to 160 receive TPT (weekly high-dose isoniazid and rifapentine for three months, 3HP) or no intervention; 161 RISK11-negative participants were randomised to no intervention or excluded from the study. In 162 the CORTIS-HR observational study, PLHIV were enrolled irrespective of RISK11 status and 163 referred for standard of care antiretroviral and isoniazid preventive therapy.

164

165 All CORTIS participants provided two spontaneously expectorated sputum samples at enrolment 166 and at the end of study visits. End of study visits were performed at month 15 of follow-up or at an 167 earlier timepoint for withdrawn participants. Both samples collected at enrolment from HIV-168 uninfected participants (CORTIS-01) were tested for Mtb using Xpert MTB/RIF, and from HIV-169 infected participants (CORTIS-HR) using Xpert MTB/RIF and MGIT culture. The two samples 170 collected at the end of study visit were tested using Xpert MTB/RIF or Xpert Ultra (Cepheid), and 171 MGIT culture. In addition, symptom-triggered TB investigations (two sputum samples; one for 172 Xpert MTB/RIF and one for MGIT culture) were performed at six routine study visits through 15-173 months follow-up. TB symptoms included at least one of: persistent unexplained cough, night 174 sweats, fever, or weight loss for 2 weeks or more, or any haemoptysis. Participants diagnosed with 175 microbiologically-confirmed TB were withdrawn from the study and referred for curative treatment. 176 All participants provided written informed consent and study protocols were approved by the 177 institutional human research ethics committees of the five participating sites.

178

#### 179 Respiratory pathobionts cohort

180 The respiratory pathobionts cohort, a subset of participants screened for eligibility for the CORTIS-01 trial at the Worcester site, has previously been described.<sup>17</sup> Pathobionts encompass pathogenic 181 182 and commensal microorganisms. Briefly, HIV-uninfected participants were consecutively enrolled 183 in this sub-study irrespective of CORTIS-01 enrolment, or signs and symptoms of upper respiratory 184 tract infections. Paired nasopharyngeal and oropharyngeal flocked swabs (FLOQSwabs, Copan 185 Diagnostics, Murrieta, CA, USA) were collected and stored in Primestore buffer (Longhorn 186 Vaccines and Diagnostics, San Antonio, TX, USA) at -80°C, and viral and bacterial nucleic acid 187 was later extracted (Qiasymphony Virus/Bacteria Mini Kit, Qiagen, Hilden, Germany) and 188 quantified using a multiplex RT-qPCR assay kit (Respiratory Pathogens 33 Kit, Fast Track 189 Diagnostics, Luxembourg) on the CFX96 Touch System lightcycler platform (Bio-Rad, Hercules, 190 CA, USA), according to the manufacturer's instructions. Participants co-enrolled into the CORTIS-191 01 trial were investigated for TB at baseline; those only enrolled into the respiratory pathobionts

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

195

#### 196 Blood sample collection and RNA extraction

197 Venous whole blood was collected in PAXgene tubes (PreAnalytiX, Hombrechtikon, Switzerland) 198 at enrolment in all cohorts, frozen at -20°C, and shipped to the South African Tuberculosis Vaccine 199 Initiative (SATVI) Cape Town laboratory. For the CTBC study, RNA was manually extracted with 200 the PAXgene blood RNA kit (Qiagen) according to the manufacturer's instructions, stored at -80 201 °C, and later used for transcriptomic analysis. For the other cohorts, RNA was extracted using a 202 high-throughput, standardised, and reproducible fully automated procedure on the Freedom EVO 203 150 robotic platform (Tecan, Männedorf, Switzerland) with the Maxwell SimplyRNA kit (Promega, 204 Madison, WI, USA). One aliquot of RNA was used immediately for cDNA synthesis and 205 measurement of the RISK11 signature, and a second aliguot was stored at -80 °C, and later used 206 to measure the panel of parsimonious transcriptomic signatures.

207

#### 208 Measurement of TB transcriptomic signatures

209 Eight parsimonious diagnostic, prognostic, and treatment response host-blood TB transcriptomic 210 signatures were pragmatically selected for inclusion in this head-to-head validation alongside 211 RISK11 based on availability of validated performance data, number of transcripts (six or fewer), 212 and availability of target sequences for custom primer-probe design or predesigned TaqMan 213 assay. As a naming convention, we use the first author and number of genes included in the signature: Francisco2<sup>23</sup>, Maertzdorf4<sup>24</sup> (also known as DIAG4), Penn-Nicholson6 (also known as 214 RISK6)<sup>22</sup>, Suliman4 (also known as RISK4)<sup>25</sup>, Roe1<sup>26</sup> (BATF2 only), Roe3<sup>27</sup>, Sweeney3<sup>28</sup>, and 215 Thompson5<sup>18</sup> (also known as RESPONSE5). The Herberg2 signature<sup>29</sup>, designed to discriminate 216 217 between viral and bacterial infection in febrile children, was also included as a potential 218 discriminator between viral and mycobacterial infection in the respiratory pathobionts cohort. 219 Signatures were measured and scores calculated as previously reported (Tables S1-S6).

220

221 Following cDNA-synthesis with EpiScript reverse transcriptase (Lucigen, Middleton, WI, USA), 222 genes of interest (Tables S2-S3) were pre-amplified using pools of TagMan primer-probe assays 223 (Thermo Fisher Scientific, Waltham, MA, USA), and gene expression (raw cycle threshold, Ct) 224 quantified by microfluidic multiplex RT-qPCR using either Fluidigm (San Francisco, CA, USA) 225 96.96 (96 samples multiplexed with 96 primer-probe assays) or 192.24 (192 samples multiplexed 226 with 24 primer-probe assays) Gene Expression chips on the BioMark HD instrument (Fluidigm). 227 Further details regarding primer-probe assay selection and gualification, data guality control, and 228 signature score measurement are provided in the **Supplemental Methods**.

230

#### 231 Endpoints

232 For the CTBC study and respiratory pathobionts sub-study, TB disease was defined by a single 233 sputum sample positive for *Mtb* on either Xpert MTB/RIF and/or liquid culture at enrolment. The 234 coprimary endpoints in the CORTIS studies were baseline prevalent TB disease and incident 235 disease through 15 months follow-up confirmed by a positive Xpert MTB/RIF, Xpert Ultra, or MGIT 236 culture, on two or more separate sputum samples collected within any 30-day period. The 237 secondary endpoint was microbiologically-confirmed TB disease on at least one sputum sample. 238 All Xpert Ultra trace positive results were excluded from the analysis because of the risk of false 239 positives.

240

#### 241 Statistical analyses

242 The statistical analysis plan for the CORTIS sub-studies was pre-specified (Appendix p20) and all 243 statistical analyses were done in R (Boston, MA, USA), version 3.6.1. Sample sizes for this sub-244 study were not pre-specified, but were based on accumulated enrolment of the CTBC and CORTIS 245 parent studies. Performance of the signatures was evaluated on their ability to diagnose prevalent 246 TB disease at baseline and predict progression to incident TB disease through 15-month follow up. 247 For all analyses, HIV-infected (CORTIS-HR) and HIV-uninfected (CORTIS-01) participants were 248 considered independently. Diagnostic performance at baseline was assessed in all participants 249 enrolled in the CORTIS studies with available PAXgene samples. Thereafter, participants meeting 250 the primary prevalent TB endpoint definition, participants who did not attend further follow-up visits, 251 and participants randomized to the RISK11-positive 3HP group (CORTIS-01 only) were excluded 252 from the primary endpoint prognostic performance analysis. Participants were followed for up to 15 253 months; those who discontinued follow-up prior to 15 months and did not meet the primary 254 endpoint definition were censored at their final study visit or last negative sputum sample collection 255 but included in the prognostic analysis.

256

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

263

For evaluation of diagnostic performance in the CTBC, CORTIS-HR, and respiratory pathobionts studies, area under the receiver operating characteristic curve (ROC) curve (AUC) was generated using the pROC package<sup>30</sup> in R. Signature performance (AUC) between HIV-infected and HIVuninfected participants in the CTBC study was compared using methods described by DeLong et

al.<sup>31</sup> Prognostic performance metrics through 15 months follow-up in CORTIS-HR were calculated by use of non-parametric methods for time-dependent ROC curve analysis from survival data using the R survAM.estimate function in the survAccuracyMeasures package<sup>32</sup> in R. Binary weighted ROC analysis was performed in the CORTIS-01 cohort to evaluate both diagnostic and prognostic performance. In a post-hoc analysis, prognostic performance in CORTIS-01 was qualitatively compared between 6, 12, and 15 months of follow-up for the primary endpoint by use of identical methods.

275

There were no pre-defined score thresholds for the parsimonious signatures; optimal signature score thresholds were calculated using the minimum Euclidean distance to 100% sensitivity and specificity, i.e. the point closest to the top-left of the ROC curve, calculated for each cut-point (*c*) as follows:  $\sqrt{[(1 - \text{Sensitivity}_{(c)})^2 + (1 - \text{Specificity}_{(c)})^2]}$ .<sup>33</sup> Signature diagnostic and prognostic accuracy was also benchmarked against the minimal WHO TPP for a community-based TB triage or referral test<sup>19</sup> (specificity 70% or sensitivity 90%) and the minimal WHO TPP for an incipient TB test, to predict progression to TB disease<sup>20</sup> (specificity 75% or sensitivity 75%), respectively.

283

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

291

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

298

#### 299 Blinding

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

- 305 microbiology results were maintained in different files, which were only integrated after the study
- 306 database had been cleaned and locked, and group allocations unblinded in CORTIS-01.
- 307
- 308 Role of the funding sources
- 309 The funders of this sub-study had no role in protocol development, study design, data collection,
- 310 data analysis, data interpretation, or writing of the report.

- 311 Results
- 312

313 Transcriptomic signatures discriminate symptomatic pulmonary TB patients from healthy,

314 asymptomatic controls

315 We translated measurement of eight parsimonious TB transcriptomic signatures to a microfluidic 316 RT- qPCR platform using TaqMan primer-probe assays for quantification of transcripts, using 317 previously described methods (Table S1-6). All new primer-probe assays passed qualification with 318 efficient linear amplification (**Table S3**). We then validated the performance of the signatures in the 319 CTBC study, a case-control cohort of HIV-uninfected and HIV-infected individuals with or without 320 pulmonary TB. Participants included 114 HIV-uninfected (53 TB cases and 61 asymptomatic 321 controls) and 86 HIV-infected (45 TB cases and 41 asymptomatic controls) adults (Figure S1A). 322 The performance of the Penn-Nicholson6 (RISK6) and RISK11 signatures in this cohort was 323 previously reported<sup>22</sup>, but is included here for comparison. We found that performance of all 324 signatures to discriminate between HIV-uninfected South African adults with TB from healthy 325 controls was excellent, with AUCs ranging from 0.89 (Thompson5, 95% CI 0.82-0.95) to 0.97 326 (RISK11 and Suliman4, both with 95% CI 0.93–1) (Figure S2). All signatures met the minimal 327 WHO triage test TPP benchmark criteria in this group (Table S7A). Discriminative performance 328 was generally lower in PLHIV, with AUCs ranging from 0.78 (Francisco2, 95% CI 0.68-0.88) to 329 0.96 (Suliman4, 95% CI 0.90–1), and was significantly worse (p<0.05) for RISK11 (p=0.027), 330 Francisco2 (p=0.0095), and Sweeney3 (p=0.021) (Figure S2). Only Penn-Nicholson6 and 331 Suliman4 met the minimal WHO triage test TPP benchmark criteria in PLHIV (Table S7B). Having 332 validated the signatures for microfluidic RT-qPCR in a case-control context, we next wished to 333 prospectively compare signature performance in a real-world setting.

334

#### 335 Recruitment of prospective community screening cohorts

In the CORTIS-01 trial,<sup>11</sup> 20,207 adults were screened at five geographically distinct community 336 337 sites throughout South Africa; 16,248 HIV-uninfected individuals met inclusion for the CORTIS-01 338 study, of whom 2,923 participants were randomized and all were included in this parsimonious 339 signature sub-study (Figure S1B and Table S8). A further 963 adults were screened and 861 340 PLHIV met inclusion for the CORTIS-HR study<sup>12</sup> and all were included in this sub-study (Figure 341 **S1D** and **Table S9**). There were 61 HIV-uninfected (weighted prevalence 1.1%, 95% CI 0.8–1.6) 342 and 10 HIV-infected (prevalence 1.2%, 95% CI 0.6-2.1) baseline primary endpoint TB cases; over 343 80% of prevalent TB cases in the HIV-uninfected cohort, and 70% in the HIV-infected cohort, were 344 asymptomatic. Thereafter, 362 HIV-uninfected participants (350 in the 3HP arm, 10 lost to follow-345 up, and 2 withdrawals) and 5 PLHIV (4 lost to follow-up and 1 withdrawal) were excluded from 346 analysis of prognostic performance for incident TB, such that 2,500 HIV-uninfected and 846 PLHIV 347 were included. Through 15 months follow-up, 24 HIV-uninfected participants (weighted incidence 348 1.1 per 100 person-years, 95% CI 0.6–1.5) and 9 PLHIV (incidence 1.0 per 100 person-years, 95% 349 CI 0.3–1.6) progressed to incident TB. Participants who discontinued follow-up prior to 15 months

- 350 were included in the prognostic assessment, but censored at their final study visit.
- 351

352 Measurement of parsimonious transcriptomic signatures in the prospective cohorts

353 The RISK11 signature assay panel (Table S2) was measured in all randomised HIV-uninfected 354 CORTIS-01 trial participants (N=2,923) and in all enrolled HIV-infected CORTIS-HR participants 355 who had baseline RNA samples available 857/861 (99.5%). RISK11 performance in these cohorts 356 was previously reported<sup>11,12</sup>, but is included here for comparison without a pre-specified score 357 threshold. The parsimonious transcriptomic signature assay panel (Table S3) was measured in 358 2,904/2,923 (99.3%) HIV-uninfected and 858/861 (99.7%) HIV-infected participants who had 359 additional baseline RNA aliquots available. The parsimonious signature failure rate ranged from 360 1.4% (RISK6) to 25.3% (Suliman4) in the HIV-uninfected cohort and from 5.2% (RISK6) to 32.1% 361 (Suliman4) in PLHIV (Tables S10-11). Failed signature results for individual samples due to failed 362 RT-qPCR reactions were not repeated, and were excluded from analysis (details in Supplemental 363 Results).

364

365 Median signature scores for RISK11, Francisco2, Maertzdorf4, RISK6, Suliman4, Sweeney3, and 366 Thompson5 at baseline were highest amongst HIV-uninfected participants with symptomatic 367 (clinical) TB, and successively lower in participants with asymptomatic (sub-clinical) and incident 368 TB (participants who progressed to TB disease), and symptomatic and asymptomatic participants 369 without TB (Figure 1). The opposite was observed for the Roe1 and Roe3 signatures; this is 370 related to data processing rather than real differences in signature biology, and a similar trend 371 would be obtained by additively inverting the scores. A similar progression was observed for most 372 signatures amongst PLHIV, although differences in the TB spectrum were not as marked, 373 especially for Francisco2 and Sweeney3 (Figure 2).

374

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

382

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

385 Diagnostic performance for primary endpoint prevalent TB was moderate for all signatures, ranging

from an AUC of 0.63 to 0.79 in the HIV-uninfected cohort (Figure 1 and Table S10). Primary

387 endpoint diagnostic performance estimates in PLHIV (AUC 0.65-0.88) were similar or exceeded 388 performance in the HIV-uninfected cohort for most signatures (Figure 2 and Table S11), except 389 the Thompson5 and Sweeney3 signatures, which appeared to perform poorly in PLHIV (lower 95%) 390 confidence bound crossing 0.50). When disaggregated by TB symptom status, diagnostic 391 performance for detection of asymptomatic (subclinical) TB among all asymptomatic participants, 392 (AUCs ranging from 0.61 to 0.78) was inferior to performance for detection of symptomatic 393 (clinical) TB among all symptomatic participants (AUCs ranging from 0.85–0.98) for all signatures 394 in HIV-uninfected participants (Figure 1). AUC estimates for detecting symptomatic TB were also 395 higher for most signatures (RISK11, Maertzdorf4, Roe3, Suliman4, and Thompson5, with AUCs 396 ranging from 0.58-0.98) amongst HIV-infected participants (Figure 2) than for detecting 397 asymptomatic, subclinical TB (AUCs ranging from 0.59-0.86). However, because only two 398 symptomatic HIV-infected and eight symptomatic HIV-uninfected participants with prevalent TB 399 had available RNA, these exploratory results should be interpreted cautiously. As the majority of 400 healthy participants and TB cases were asymptomatic, the AUC estimates for the pooled primary 401 endpoint are very similar to estimates for the asymptomatic subgroup. Diagnostic performance for 402 the secondary endpoint was not qualitatively different to that of the primary endpoint for all 403 signatures in both HIV-uninfected and HIV-infected cohorts (Figure S3-4).

404

405 For the primary diagnostic endpoint and asymptomatic sub-group, no signatures met the minimal 406 WHO TPP benchmarks for a triage test (sensitivity 90% and specificity 70%) in the HIV-uninfected 407 cohort (Figures 1 and Table S10). However, five signatures (RISK11, RISK6, Roe1, Roe3 and 408 Suliman4) met the optimal WHO TPP benchmark (sensitivity 95% and specificity 80%) amongst 409 symptomatic participants, and the other four signatures (Francisco2, Maertzdorf4, Sweeney3, and 410 Thompson5) either approached or met the minimal WHO TPP in this sub-group (Figure 1). 411 Similarly amongst PLHIV, no signatures met the minimal WHO TPP benchmarks for a triage test 412 (sensitivity 90% and specificity 70%) for the primary diagnostic endpoint and asymptomatic sub-413 group (Figure 2 and Table S11), however four signatures (RISK11, Maertzdorf4, Suliman4, and 414 Thompson5) met the optimal WHO triage test benchmarks in participants with TB symptoms 415 (Figure 2). Due to the small numbers of primary endpoint prevalent TB cases with evaluable 416 transcriptomic signature scores amongst PLHIV (two symptomatic and six asymptomatic), 417 precision around estimates was limited. For the primary diagnostic endpoint in the HIV-infected 418 cohort, the upper limits of the 95% confidence intervals of performance metrics were within the 419 bounds of the minimal WHO triage test criteria for most signatures, barring Francisco2, Sweeney3, 420 and Thompson5 (Table S11).

421

422 Significant prognostic performance of signatures through 12 months in HIV-uninfected and 15-423 months in HIV-infected cohorts 424 We next evaluated transcriptomic signature prognostic performance for incident TB in all 425 participants without primary endpoint prevalent TB at baseline who attended one or more follow-up 426 visit. All signatures measured at enrolment had excellent performance for detecting HIV-uninfected 427 participants who progressed to primary endpoint incident TB disease through six months follow-up, 428 with AUCs ranging from 0.80–0.95 (Figure 3 and Table S12). Prognostic performance declined 429 through 15 months follow-up (AUCs 0.49–0.66) with only RISK6 showing significant discrimination 430 at this timepoint (AUC 0.66, 95% CI 0.51–0.82). Most signatures demonstrated good discrimination 431 (AUCs 0.63–0.80) up to 12 months post enrolment. In the HIV-infected CORTIS-HR cohort, there 432 were insufficient primary endpoint incident TB cases to evaluate prognostic performance prior to 433 15 months follow-up. Six of nine transcriptomic signatures were able to significantly differentiate 434 between incident TB cases and controls in the HIV-infected cohort through 15 months, with AUCs 435 ranging from 0.54–0.81 (Figures 4 and Tables S13). Lower 95% confidence bounds for the 436 Francisco2, Sweeney3, and Thompson5 signatures crossed 0.50 indicating that prognostic 437 performance at this timepoint was not significant. Prognostic performance through 15 months 438 follow-up for the secondary endpoint was not qualitatively different to the primary endpoint in either 439 cohort (Figure S5-6).

440

441 For primary endpoint incident TB, none of the signatures met the minimal WHO TPP benchmark 442 for a prognostic test (sensitivity 75% and specificity 75%) in the HIV-uninfected cohort through 15 443 months follow-up (Figure 3 and Table S12); the Roe3 signature came closest at its optimal cut-444 point (sensitivity 70.0%, 95% CI 46.0-92.9; specificity 72.9%, 95% CI 71.1-74.7). However, through 445 six months follow-up, six of eight parsimonious signatures met the minimal benchmark. The 446 remaining two signatures, Suliman4 and Sweeney3, approached it with upper limits of the 95% 447 confidence intervals for prognostic estimates exceeding the benchmark bounds (Table S12). Only 448 RISK11 met the optimal WHO TPP benchmark for a prognostic test (sensitivity 90% and specificity 449 90%) through six months. Through 12 months follow-up, the upper limits of the 95% confidence 450 intervals for all signatures exceeded the minimal WHO TPP benchmark, but only the Roe3 451 signature (sensitivity 84.9%, 95% CI 55.0-100; specificity 75.4%, 95% CI 73.7-77.1) met the 452 benchmark. Amongst people living with HIV, only the Roe1 signature (sensitivity 83.8%, 95% CI 453 28.1-98.6; specificity 78.0%, 95% CI 74.5-81.0) met the minimal WHO TPP benchmark for a 454 prognostic test through 15 months follow-up (Figure 4 and Table S13), however the upper 95% 455 confidence interval bounds for six of eight of the other signatures (except Maertzdorf4 and 456 Sweeney3) also met or approximated this benchmark.

457

458 Most TB signatures failed to differentiate participants with upper respiratory viral pathobionts from 459 those with TB disease

460 Amongst participants with HIV we found that those without TB who had a detectable HIV plasma 461 viral load (≥100 copies/mL) had significantly higher (p<0.0001; or conversely lower for Roe1 and Roe3) transcriptomic signature scores than those with supressed HIV plasma viral load (<100 copies/mL). The exception was Thompson5, the only signature that does not include ISGs<sup>18</sup>, which was not affected by viral load (p=0.87; **Figure S7**). Among HIV-uninfected participants without TB, scores for most signatures (except Francisco2 and Sweeney3) differed significantly (p<0.05) between those with and without TB symptoms (**Figure 1**); we hypothesised that these differences were due to the presence of other comorbid infections in the symptomatic participants.

468

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

476

477 TB transcriptomic signature scores were significantly higher (p<0.05; or conversely lower for Roe1 478 and Roe3) in participants with any viral pathobiont, or both viral and bacterial pathobionts, 479 compared to those with only bacterial, or no pathobionts, with the exception of Thompson5 (Figure 480 5). Most of the signatures significantly discriminated between viral and bacterial pathobionts, and 481 viral and no pathobionts, but were unable to differentiate bacterial from no pathobionts. We also measured the Herberg2 signature<sup>29</sup>, designed to discriminate between viral and bacterial infection 482 483 in febrile children, in the respiratory pathobionts cohort; this signature performed equivalently to the 484 other signatures in differentiating between the groups. Most of the upper respiratory bacteria 485 detected (H influenza 21.3%, S aureus 10.2%, S pneumonia 9.2%, M catarrhalis 5.8%) commonly 486 colonise the respiratory tract of healthy individuals, hence we are unable to determine the effect of 487 pathogenic bacteria on signature scores. Thompson5 was the only signature for which no 488 differences in scores were observed between participants with viral, bacterial, or no pathobionts, 489 and subsequently was not able to differentiate between these groups (Figure 5).

490

491 28.6% (286/1,000) of participants were co-enrolled into the respiratory pathobionts sub-study and 492 CORTIS-01 and consequently investigated for TB (Figure S1C). There were insufficient primary 493 endpoint TB cases for meaningful analysis. However, 11 secondary endpoint (≥1 sputum sample 494 positive for *Mtb*) prevalent TB cases (3.8%) were identified in CORTIS-01; 4/99 (4.0%) participants 495 with bacteria detected in the upper respiratory tract and 7/150 (4.7%) with no pathobionts detected. 496 No participants with viral upper respiratory pathobionts only (n=41), or mixed viral and bacterial 497 pathobionts (n=33), had prevalent TB. The presence of bacterial upper respiratory pathobionts did 498 not significantly affect TB diagnostic performance. Transcriptomic signatures differentiated 499 between participants with prevalent TB and those with no pathobionts detected, with AUCs

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

#### 504 **Discussion**

505 We validated eight parsimonious TB transcriptomic signatures by microfluidic RT-qPCR and 506 prospectively evaluated their diagnostic and prognostic accuracy in two large community screening 507 cohorts of unselected HIV-uninfected and HIV-infected adults. We found that none of the signatures met the WHO triage test TPP benchmarks<sup>19</sup> for subclinical TB, although performance 508 509 for several signatures approached these thresholds. Parsimonious signature diagnostic 510 performance in adults with TB symptoms was superior to that in asymptomatic participants, with 511 most meeting or approaching the TPP benchmarks. In the HIV-uninfected cohort prognostic 512 accuracy for incident TB within six months of testing was excellent for virtually all parsimonious 513 signatures, while performance for TB within 12 months was moderate. Only one signature, RISK6, 514 had any statistically significant prognostic ability for incident TB within 15 months of testing.

515

516 This waning accuracy mirrors the finding of a short prognostic window after signature testing in the 517 meta-analysis by Gupta et al<sup>14</sup>, and supports the hypothesis that the timeframe of TB progression 518 from infection to clinical disease is highly heterogenous between individuals, ranging from months 519 to vears.<sup>36-38</sup> We postulate that the declining performance distal to signature measurement may be 520 due to low inflammatory activity in early disease stages, described as incipient TB. Individuals with 521 incipient TB may spontaneously control the infection and thus not progress to disease<sup>37</sup>, resulting 522 in reduced specificity. In this regard, we suggest that the 2-year predictive window for a prognostic 523 test proposed in the WHO framework<sup>20</sup> be revisited. Transient viral infections may also increase 524 false positivity.<sup>17</sup> Low sensitivity may reflect the high burden of population *Mtb* exposure and rate of 525 new infections in this endemic setting.

526

527 Performance of parsimonious signatures in PLHIV appeared to match or surpass performance in 528 participants without HIV, consistent with the findings of Turner and colleagues<sup>15</sup>. Most signatures were able to differentiate between those who would progress to TB disease (progressors) through 529 530 15 months follow-up and non-progressors in the HIV-infected cohort, and the Roe1 signature met the minimal WHO TPP benchmark for a prognostic test<sup>20</sup>. The greater prognostic window of the 531 532 signatures measured at baseline in PLHIV may reflect prevalent TB cases missed at study 533 enrolment (possibly due to paucibacillary sputum), but detected further along the disease spectrum 534 during follow-up. At each study visit, participants underwent symptom screening with symptom-535 guided TB investigations which would have missed asymptomatic disease at earlier timepoints. 536 However, all participants provided sputum samples for TB investigation at the end-of-study visits, 537 which may also explain the bulk of disease diagnosed at the month 15 visit. Most of the incident 538 TB cases in the HIV-infected cohort were detected after 12 months follow-up; this may be due to 539 the appearance of symptoms late in the disease course in immunosuppressed individuals, and as 540 a result, incident TB cases may have been missed by symptom screening at earlier visits. It is also 541 conceivable that the better performance in PLHIV is driven by ART-naïve individuals who have increased risk of TB disease, but also have elevated signature scores (or lower for Roe1 and Roe3) due to their higher HIV plasma viral loads. However, 74% (142/193) of ART-naïve participants started ART during the study (**Table S9**), which would lower their risk of progression, and increase "false" positivity. Unfortunately our study was underpowered to evaluate signature performance stratified by viral load or ART status, however we postulate that higher signatures scores in those with non-supressed HIV plasma viral loads (**Figure S7**) would result in significantly lower specificity, but equivalent or higher sensitivity.

549

550 We recently showed that elevated RISK11 signature scores in healthy individuals without TB return 551 to normal levels within three months irrespective of the use of TPT in the majority of individuals.<sup>17</sup> 552 The transient elevation of RISK11 may be due to self-clearance of *Mtb* infection in certain 553 individuals, but is likely due to other exogenous factors that temporarily affect transcriptomic 554 signature gene expression. Scores and performance of most parsimonious signatures, with the 555 exception of the Thompson5 signature, were affected by the presence of upper respiratory viral, 556 but not bacterial, pathobionts. Most transcriptomic signatures are primarily comprised of ISGs, 557 which are known to be induced by viral infections, including HIV. The Thompson5 signature, which 558 does not contain any ISGs, was unaffected by respiratory pathobionts or HIV. This supports the 559 hypothesis that exogenous factors, and in particular viral infections, have a large impact on TB 560 transcriptomic signature scores and performance due to induction of these genes; design of transcriptomic signatures which do not predominantly consist of ISGs, 39,40 or which use a 561 multinomial modelling<sup>41</sup> approach to exclude viral infection, may address this shortcoming. 562

563

564 Given that this study was not designed to compare diagnostic and prognostic accuracy between 565 signatures, and the 95% confidence intervals of performance metrics for most signatures were 566 overlapping, we are hesitant to champion any particular signature. Several parsimonious 567 signatures performed at least as well as RISK11. Performance of most signatures is promising for 568 triage of symptomatic adults seeking care and PLHIV attending ART clinic, to rule out active 569 disease and guide further investigation, and for short-term (6-12 months) prognosis of TB disease 570 risk for targeted preventive therapy. Recent studies of 3HP have shown limited duration of efficacy 571 (9 months) for biomarker-guided preventive therapy in HIV-uninfected adults in South Africa<sup>11</sup>, and 572 no additional benefit of universal annual retreatment of PLHIV on ART compared to a single 573 course<sup>42</sup>. Our results further suggest that a prognostic window greater than 12 months after testing 574 is unlikely to be of value in high transmission settings with associated risk of re-infection. This 575 implies that annual re-testing would likely be necessary if such signatures are to be included in TB control strategies.<sup>43</sup> Parsimonious signature performance reported here warrants translation to 576 577 point-of-care devices for clinical implementation studies.

579 No signatures met the WHO triage test TPP accuracy benchmark for diagnosing asymptomatic 580 prevalent TB, which is a setback for community active case-finding efforts in endemic settings, such as South Africa, where there is a high burden of undiagnosed sub-clinical disease.<sup>2-5,11,12</sup> 581 582 However, we would note that these benchmarks were not intended for TB screening in a healthy 583 community setting, but rather "to test persons with any symptoms or risk factors suggestive of TB 584 who are seeking care"<sup>19</sup>, and thus the bar may have been too high for a strategy that seeks to find subclinical TB cases. Currently, chest radiography is the only sufficiently accurate screening tool in 585 widespread use for detecting asymptomatic sub-clinical disease.<sup>2-5</sup> Portable molecular screening 586 587 has also been proposed as a possible universal screening tool for active case-finding,<sup>7</sup> but this 588 remains unaffordable in most settings. Given the recently recognised burden of asymptomatic TB 589 in endemic settings and lack of affordable and accessible active case-finding tools to detect early 590 and subclinical disease in community settings, we would propose that a new consensus framework 591 for the development of screening tests for asymptomatic TB is urgently needed. Such a framework 592 should also provide clear definitions of contentious concepts of asymptomatic "incipient" and "subclinical" TB.<sup>44,45</sup> We have previously demonstrated a sequential inflammatory TB disease process 593 594 <sup>22,36,46</sup>. The trend of increasing signature scores from individuals without TB disease, through early 595 or incipient (individuals who later progress to incident disease), sub-clinical, and clinical TB 596 disease observed in this study (Figures 1-2) succinctly illustrates the stages of disease 597 progression. Further trials are needed to determine whether earlier stages of TB progression 598 require a full course of curative treatment, or an abbreviated regimen such as 3HP.

599

600 The correlation between most signatures (with the exception of Thompson5), and fact that no 601 individual signature was significantly superior for diagnosis or prognosis, indicates they likely 602 measure the same or similar biological pathways. Our study would suggest that performance of 603 transcriptomic host-response signatures, based on discovery in active TB cases or progressor 604 cohorts, has reached an apex and that further signature discovery using similar conventional 605 methodology and study designs is unlikely to deliver significantly better diagnostic tools. What 606 could be done differently? First, we would argue the necessity for recruiting larger heterogenous 607 unselected prospective discovery cohorts, with TB cases representing a spectrum from sub-clinical 608 through clinical disease at prevalence rates that reflect the real epidemic, rather than carefully-609 curated case-control studies with uniform and highly enriched TB cases and homogenous healthy 610 controls or inappropriate other diseases. In this study, all signatures performed superbly in the 611 CTBC case-control cohort, but performance dipped substantially in the real-world CORTIS studies. 612 Second, we have shown that parsimonious signatures perform at least as well as non-reduced 613 signatures such as RISK11; early translation of parsimonious signatures to near point-of-care 614 technology, such as RT-qPCR, would allow more rapid advancement along the developmental 615 pipeline. Third, few studies have incorporated clinical features (such as age, sex, body-mass 616 index, C-reactive protein, haemoglobin, monocyte-lymphocyte ratio, and HIV status or viral load)

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

623

624 A major strength of this study was the prospective recruitment of large community cohorts of 625 unselected adults from multiple sites reflecting distinct patterns of *Mtb* exposure and rates of TB 626 disease in an endemic setting. Despite intensive community active case-finding with enrichment of 627 the HIV-uninfected study population for individuals at greatest risk of TB disease, sputum sample 628 collection from all enrolled participants at enrolment and end-of-study visits, and symptom-629 triggered investigation during follow-up, there were relatively few prevalent and incident TB cases, 630 limiting power for analysis of diagnostic and prognostic performance. We also did not investigate 631 participants for extra-pulmonary TB, potentially missing cases amongst PLHIV. Another limitation 632 was the high Suliman4 failure rate, resulting in less precise performance estimates for this 633 signature. We also acknowledge that primer sequences selected for RT-qPCR may not be 634 equivalent to those used for microarray or in the Cepheid Xpert MTB Host Response prototype cartridge<sup>49,50</sup>. The large proportion of asymptomatic prevalent TB cases is a feature unique to this 635 636 cohort, and may prove valuable for future biomarker discovery. Recruitment was restricted to 637 South Africa, limiting generalisability of the results to other contexts; there is a need for similar 638 large prospective diagnostic validation cohorts in other TB endemic regions.

639

640 Several parsimonious TB transcriptomic signatures hold promise for triage of symptomatic adults 641 seeking care, screening of ART clinic attendees and other high-risk groups for further investigation, 642 and prediction of short-term risk of TB for initiation of targeted preventive therapy. Point-of-care 643 RT-qPCR testing platforms currently under development and in field implementation studies,<sup>49,50</sup> 644 would bring these signatures from the laboratory bench to the clinic bedside.

#### 645 Author contributions

646 SCM, AP-N, AF-G, MH, and TJS conceived the study. AP-N, SKM, MH, and TJS implemented 647 clinical studies, raised funds and/or provided the resources. MT, GW, KN, and GC were 648 responsible for all site-level activities, including recruitment, clinical management, and data 649 collection. SKM, MM, HM, MF, KH, and ME, and provided operational, technical, or laboratory 650 support and project management. SCM, ME, and ON processed samples and performed the 651 experiments. SCM, SKM, AF-G, and HM verified the underlying data. SCM and AF-G analysed 652 data. SCM, SKM, TJS, and MH interpreted results and wrote the first draft of the manuscript. 653 Members of the CTBC Study Team and the CORTIS Study Team performed clinical studies and 654 processed samples. All authors had full access to the data, confirm the integrity of the data and its 655 presentation, agree with its interpretation as discussed in the manuscript, and reviewed, revised, 656 and approved the manuscript before submission. The corresponding author had final responsibility 657 for the decision to submit for publication.

658

#### 659 **Competing interests statement**

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

667

#### 668 Data availability

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

674

#### 675 Funding

The CTBC study was funded by the Strategic Health Innovation Partnerships (SHIP) Unit of the South African Medical Research Council (SAMRC) with funds received from the South African Department of Science and Technology. The CORTIS studies were funded by the Bill and Melinda Gates Foundation (BMGF; OPP1116632, OPP1137034, and OPP1151915) and also the SHIP Unit of the SAMRC with funds received from the South African Department of Science and Technology. Respiratory infection assays were funded by the NIH (Al123780).

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

- 691 References
- 692
- 693 1. WHO. Global tuberculosis report 2020. Geneva: World Health Organization, 2020. 694 https://www.who.int/tb/publications/global report/en/ (accessed October 15, 2020). 695 2. Onozaki I, Law I, Sismanidis C, Zignol M, Glaziou P, Floyd K. National tuberculosis 696 prevalence surveys in Asia, 1990-2012: an overview of results and lessons learned. 697 Tropical Medicine and International Health. 2015;20(9):1128-45. doi: 10.1111/tmi.12534. 698 3. Frascella B, Richards AS, Sossen B, Emery JC, Odone A, Law I, et al. Subclinical 699 tuberculosis disease - a review and analysis of prevalence surveys to inform definitions, 700 burden, associations and screening methodology. Clinical Infectious Diseases. 2020. doi: 701 10.1093/cid/ciaa1402. 702 4. Law I, Floyd K, African TBPSG. National tuberculosis prevalence surveys in Africa, 2008-703 2016: an overview of results and lessons learned. Tropical Medicine and International 704 Health. 2020;25(11):1308-27. doi: 10.1111/tmi.13485. 705 5. South African National Department of Health. The First National TB Prevalence Survey: 706 South Africa 2018. 2021. www.knowledgehub.org.za/elibrary/first-national-tb-prevalence-707 survey-south-africa-2018 (accessed May 22, 2021). 708 6. Harris M, Qi A, Jeagal L, Torabi N, Menzies D, Korobitsyn A, et al. A systematic review of 709 the diagnostic accuracy of artificial intelligence-based computer programs to analyze chest 710 x-rays for pulmonary tuberculosis. PLoS ONE. 2019;14(9):e0221339. doi: 711 10.1371/journal.pone.0221339. 712 7. Calligaro GL, Zijenah LS, Peter JG, Theron G, Buser V, McNerney R, et al. Effect of new 713 tuberculosis diagnostic technologies on community-based intensified case finding: a 714 multicentre randomised controlled trial. Lancet Infectious Diseases. 2017;17(4):441-50. doi: 715 10.1016/S1473-3099(16)30384-X. 716 8. Fiore-Gartland A, Carpp LN, Naidoo K, Thompson E, Zak DE, Self S, et al. Considerations 717 for biomarker-targeted intervention strategies for tuberculosis disease prevention. 718 Tuberculosis (Edinb). 2018;109:61-8. doi: 10.1016/j.tube.2017.11.009. 719 9. Darboe F, Mbandi SK, Thompson EG, Fisher M, Rodo M, van Rooyen M, et al. Diagnostic 720 performance of an optimized transcriptomic signature of risk of tuberculosis in 721 cryopreserved peripheral blood mononuclear cells. Tuberculosis (Edinb). 2018;108:124-6. 722 doi: 10.1016/j.tube.2017.11.001. 723 Darboe F, Mbandi SK, Naidoo K, Yende-Zuma N, Lewis L, Thompson EG, et al. Detection 10. 724 of Tuberculosis Recurrence, Diagnosis and Treatment Response by a Blood 725 Transcriptomic Risk Signature in HIV-Infected Persons on Antiretroviral Therapy. Front 726 Microbiol. 2019;10:1441. doi: 10.3389/fmicb.2019.01441.

727 11. Scriba TJ, Fiore-Gartland A, Penn-Nicholson A, Mulenga H, Kimbung Mbandi S, Borate B, 728 et al. Biomarker-guided tuberculosis preventive therapy (CORTIS): a randomised controlled 729 trial. Lancet Infectious Diseases. 2021;21(3):354-65. doi: 10.1016/S1473-3099(20)30914-2. 730 12. Mendelsohn SC, Fiore-Gartland A, Penn-Nicholson A, Mulenga H, Mbandi SK, Borate B, et 731 al. Validation of a host blood transcriptomic biomarker for pulmonary tuberculosis in people 732 living with HIV: a prospective diagnostic and prognostic accuracy study. Lancet glob health. 733 2021;9(6):e841-e53. doi: 10.1016/S2214-109X(21)00045-0. 734 Warsinske H, Vashisht R, Khatri P. Host-response-based gene signatures for tuberculosis 13. 735 diagnosis: A systematic comparison of 16 signatures. PLoS Medicine. 736 2019;16(4):e1002786. doi: 10.1371/journal.pmed.1002786. 737 14. Gupta RK, Turner CT, Venturini C, Esmail H, Rangaka MX, Copas A, et al. Concise whole 738 blood transcriptional signatures for incipient tuberculosis: a systematic review and patient-739 level pooled meta-analysis. Lancet Respir Med. 2020;8(4):395-406. doi: 10.1016/S2213-740 2600(19)30282-6. 741 Turner CT, Gupta RK, Tsaliki E, Roe JK, Mondal P, Nyawo GR, et al. Blood transcriptional 15. 742 biomarkers for active pulmonary tuberculosis in a high-burden setting: a prospective, 743 observational, diagnostic accuracy study. Lancet Respir Med. 2020;8(4):407-19. doi: 744 10.1016/S2213-2600(19)30469-2. 745 Mulenga H, Zauchenberger CZ, Bunyasi EW, Mbandi SK, Mendelsohn SC, Kagina B, et al. 16. 746 Performance of diagnostic and predictive host blood transcriptomic signatures for 747 Tuberculosis disease: A systematic review and meta-analysis. PLoS ONE. 748 2020;15(8):e0237574. doi: 10.1371/journal.pone.0237574. 749 17. Mulenga H, Musvosvi M, Mendelsohn SC, Penn-Nicholson A, Mbandi SK, Fiore-Gartland 750 A, et al. Longitudinal dynamics of a blood transcriptomic signature of tuberculosis. 2021 751 (Submitted). 752 18. Thompson EG, Du Y, Malherbe ST, Shankar S, Braun J, Valvo J, et al. Host blood RNA 753 signatures predict the outcome of tuberculosis treatment. Tuberculosis (Edinb). 754 2017;107:48-58. doi: 10.1016/j.tube.2017.08.004. 755 19. WHO. High-priority target product profiles for new tuberculosis diagnostics: report of a 756 consensus meeting. Geneva: World Health Organization, 2014. 757 www.who.int/tb/publications/tpp report/en/ (accessed May 22, 2020). 758 20. WHO. Consensus Meeting Report: Development of a Target Product Profile (TPP) and a 759 framework for evaluation for a test for predicting progression from tuberculosis infection to 760 active disease. Geneva: World Health Organization, 2017. 761 http://apps.who.int/iris/handle/10665/259176 (accessed May 22, 2020). 762 21. Bossuvt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig L, et al. STARD 763 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 764 2015;351:h5527. doi: 10.1136/bmj.h5527.

765 22. Penn-Nicholson A, Mbandi SK, Thompson E, Mendelsohn SC, Suliman S, Chegou NN, et 766 al. RISK6, a 6-gene transcriptomic signature of TB disease risk, diagnosis and treatment 767 response. Sci Rep. 2020;10(1):8629. doi: 10.1038/s41598-020-65043-8. 768 23. Francisco NM, Fang YM, Ding L, Feng S, Yang Y, Wu M, et al. Diagnostic accuracy of a 769 selected signature gene set that discriminates active pulmonary tuberculosis and other 770 pulmonary diseases. Journal of Infection. 2017;75(6):499-510. doi: 771 10.1016/j.jinf.2017.09.012. 772 Maertzdorf J, McEwen G, Weiner J, 3rd, Tian S, Lader E, Schriek U, et al. Concise gene 24. 773 signature for point-of-care classification of tuberculosis. EMBO Molecular Medicine. 774 2016;8(2):86-95. doi: 10.15252/emmm.201505790. 775 25. Suliman S, Thompson EG, Sutherland J, Weiner J, 3rd, Ota MOC, Shankar S, et al. Four-776 Gene Pan-African Blood Signature Predicts Progression to Tuberculosis. American Journal 777 of Respiratory and Critical Care Medicine. 2018;197(9):1198-208. doi: 778 10.1164/rccm.201711-2340OC. 779 Roe JK, Thomas N, Gil E, Best K, Tsaliki E, Morris-Jones S, et al. Blood transcriptomic 26. 780 diagnosis of pulmonary and extrapulmonary tuberculosis. JCI Insight. 2016;1(16):e87238. 781 doi: 10.1172/jci.insight.87238. 782 27. Roe J, Venturini C, Gupta RK, Gurry C, Chain BM, Sun Y, et al. Blood Transcriptomic 783 Stratification of Short-term Risk in Contacts of Tuberculosis. Clinical Infectious Diseases. 784 2020;70(5):731-7. doi: 10.1093/cid/ciz252. 785 Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of 28. 786 pulmonary tuberculosis: a multicohort analysis. Lancet Respir Med. 2016;4(3):213-24. doi: 787 10.1016/S2213-2600(16)00048-5. 788 29. Herberg JA, Kaforou M, Wright VJ, Shailes H, Eleftherohorinou H, Hoggart CJ, et al. 789 Diagnostic Test Accuracy of a 2-Transcript Host RNA Signature for Discriminating Bacterial 790 vs Viral Infection in Febrile Children. JAMA. 2016;316(8):835-45. doi: 791 10.1001/jama.2016.11236. 792 30. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, et al. pROC: an open-793 source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 794 2011;12:77. doi: 10.1186/1471-2105-12-77. 795 31. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more 796 correlated receiver operating characteristic curves: a nonparametric approach. Biometrics. 797 1988;44(3):837-45. doi: 10.2307/2531595. 798 32. Zheng Y, Cai T, Brown M. Package 'survAccuracyMeasures'. 2014. 799 http://mdbrown.github.io/survAccuracyMeasures/ (accessed May 22, 2020). 800 33. Perkins NJ, Schisterman EF. The inconsistency of "optimal" cutpoints obtained using two 801 criteria based on the receiver operating characteristic curve. American Journal of 802 Epidemiology. 2006;163(7):670-5. doi: 10.1093/aje/kwj063.

803 34. Carpenter J, Bithell J. Bootstrap confidence intervals: when, which, what? A practical guide 804 for medical statisticians. Statistics in Medicine. 2000;19(9):1141-64. doi: 805 10.1002/(sici)1097-0258(20000515)19:9<1141::Aid-sim479>3.0.Co;2-f. 806 35. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful 807 approach to multiple testing. Journal of the Royal statistical society: series B 808 (Methodological). 1995;57(1):289-300. 809 36. Scriba TJ, Penn-Nicholson A, Shankar S, Hraha T, Thompson EG, Sterling D, et al. 810 Sequential inflammatory processes define human progression from M. tuberculosis 811 infection to tuberculosis disease. PLoS Pathogens. 2017;13(11):e1006687. doi: 812 10.1371/journal.ppat.1006687. 813 37. Drain PK, Bajema KL, Dowdy D, Dheda K, Naidoo K, Schumacher SG, et al. Incipient and 814 Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. 815 Clinical Microbiology Reviews. 2018;31(4):e00021-18. doi: 10.1128/CMR.00021-18. 816 38. Scriba TJ, Mendelsohn SC. Headway made towards biosignatures for incipient 817 tuberculosis. Lancet Respir Med. 2020;8(4):328-30. doi: 10.1016/S2213-2600(19)30355-8. 818 39. Singhania A, Verma R, Graham CM, Lee J, Tran T, Richardson M, et al. A modular 819 transcriptional signature identifies phenotypic heterogeneity of human tuberculosis 820 infection. Nature Communications. 2018;9(1):2308. doi: 10.1038/s41467-018-04579-w. 821 40. Esmail H, Lai RP, Lesosky M, Wilkinson KA, Graham CM, Horswell S, et al. Complement 822 pathway gene activation and rising circulating immune complexes characterize early 823 disease in HIV-associated tuberculosis. Proceedings of the National Academy of Sciences 824 of the United States of America. 2018;115(5):E964-E73. doi: 10.1073/pnas.1711853115. 825 41. Duffy FJ, Thompson EG, Scriba TJ, Zak DE. Multinomial modelling of TB/HIV co-infection 826 yields a robust predictive signature and generates hypotheses about the HIV+TB+ disease 827 state. PLoS ONE. 2019;14(7):e0219322. doi: 10.1371/journal.pone.0219322. 828 42. Churchyard G, Cardenas V, Chihota V, Mngadi K, Sebe M, Brumskine W, et al. 829 Effectiveness of 3HP annually vs once for HIV-positive people: The WHIP3TB Trial [CRO] 830 Abstract 143]. In Special Issue: Abstracts From the 2020 Conference on Retroviruses and 831 Opportunistic Infections. Top Antivir Med. 2020;28(1). 832 43. Sumner T, Mendelsohn SC, Scriba TJ, Hatherill M, White R. The impact of blood 833 transcriptomic biomarker targeted tuberculosis preventive therapy in people living with HIV: 834 a mathematical modelling study. 2021 (Submitted). 835 44. Kendall EA, Shrestha S, Dowdy DW. The Epidemiological Importance of Subclinical 836 Tuberculosis. A Critical Reappraisal. American Journal of Respiratory and Critical Care 837 Medicine. 2021;203(2):168-74. doi: 10.1164/rccm.202006-2394PP. 838 45. Pierce CW. Subclinical Tuberculosis: Some Flies in the Ointment. American Journal of 839 Respiratory and Critical Care Medicine. 2021. doi: 10.1164/rccm.202101-0227LE.

840 46. Zak DE, Penn-Nicholson A, Scriba TJ, Thompson E, Suliman S, Amon LM, et al. A blood 841 RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet. 842 2016;387(10035):2312-22. doi: 10.1016/S0140-6736(15)01316-1. 843 47. Sivakumaran D, Jenum S, Vaz M, Selvam S, Ottenhoff THM, Haks MC, et al. Combining 844 host-derived biomarkers with patient characteristics improves signature performance in 845 predicting tuberculosis treatment outcomes. Commun Biol. 2020;3(1):359. doi: 846 10.1038/s42003-020-1087-x. 847 48. Mendelsohn SC, Mbandi SK, Hatherill M, Scriba TJ. Blood transcriptional signatures for 848 tuberculosis testing. Lancet Respir Med. 2020;8(4):330-1. doi: 10.1016/S2213-849 2600(20)30045-X. 850 49. Sodersten E, Ongarello S, Mantsoki A, Wyss R, Persing DH, Banderby S, et al. Diagnostic 851 Accuracy Study of a Novel Blood-Based Assay for Identification of Tuberculosis in People 852 Living with HIV. Journal of Clinical Microbiology. 2021;59(3). doi: 10.1128/JCM.01643-20. 853 50. Moreira FMF, Verma R, Pereira Dos Santos PC, Leite A, da Silva Santos A, de Araujo 854 RCP, et al. Blood-based host biomarker diagnostics in active case finding for pulmonary 855 tuberculosis: A diagnostic case-control study. EClinicalMedicine. 2021;33:100776. doi: 856 10.1016/j.eclinm.2021.100776. 857 858

859 Figure legends

860

- Figure 1: Parsimonious signature diagnostic performance for prevalent TB and signature
   score distributions in people without HIV
- 863 Paired receiver operating characteristic (ROC) curves and box-and-whisker plots for the (A-B)
- 864 Darboe11 (RISK11), (C-D) Francisco2, (E-F) Maertzdorf4 (DIAG4), (G-H) Penn-Nicholson6
- 865 (RISK6), (I-J) Roe1 (BATF2), (K-L) Roe3, (M-N) Suliman4 (RISK4), (O-P) Sweeney3, and (Q-R)
- 866 Thompson5 (RESPONSE5) signatures in the CORTIS-01 study of people without HIV.
- 867 The ROC curves depict diagnostic performance (area under the curve, AUC, with 95% CI) of the
- 868 parsimonious signatures for the primary endpoint (1' Endpoint), i.e. TB diagnosed on two or more
- 869 liquid culture-positive or Xpert MTB/RIF-positive sputum samples. The ROC curves show
- 870 participants with symptomatic clinical prevalent TB versus symptomatic controls (Symptom+), and
- 871 participants with asymptomatic, subclinical prevalent TB versus asymptomatic controls (Symptom-
- 872 ). The shaded areas represent the 95% CIs. The solid box depicts the optimal criteria (95%
- 873 sensitivity and 80% specificity) and the dashed box depicts the minimal criteria (90% sensitivity
- and 70% specificity) set out in the WHO Target Product Profile for a triage test.<sup>19</sup>
- 875 The box-and-whisker plots depict signature score (measured at enrolment) distribution by
- 876 symptom status (each dot represents a participant) in asymptomatic and symptomatic participants
- 877 with no TB, participants who progressed to incident TB, and participants with prevalent subclinical
- 878 (asymptomatic) and clinical (symptomatic) TB. Prevalent and incident TB comprised all primary
- 879 endpoint cases. Symptoms were recorded at the time of enrolment for participants without TB and
- 880 those with prevalent TB. p values for comparison of median signature scores between groups in
- 881 box-and-whisker plots were calculated with the Mann-Whitney U test and corrected for multiple
- comparisons by use of the Benjamini-Hochberg Procedure.<sup>35</sup> Boxes depict the IQR, the midline
- represents the median, and the whiskers indicate the IQR  $\pm$  (1.5 × IQR).
- 884 (S) Summary of signature diagnostic performance in order of primary endpoint AUC estimates.
- 885 The diagnostic AUC estimates in symptomatic and asymptomatic participant sub-groups are also
- shown. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, and the black
- 887 dotted line indicates the lower bound of the 95% CI for the best performing signature for the
- 888 primary endpoint.
- 889

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

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

896 The ROC curves depict diagnostic performance (area under the curve, AUC, with 95% CI) of the 897 parsimonious signatures for TB diagnosed for the primary endpoint (1' Endpoint), i.e. TB 898 diagnosed on two or more liquid culture-positive or Xpert MTB/RIF-positive sputum samples. The 899 ROC curves show participants with symptomatic clinical prevalent TB versus symptomatic controls 900 (Symptom+), and participants with asymptomatic subclinical prevalent TB versus asymptomatic 901 controls (Symptom-). The shaded areas represent the 95% CIs. The solid box depicts the optimal 902 criteria (95% sensitivity and 80% specificity) and the dashed box depicts the minimal criteria (90% 903 sensitivity and 70% specificity) set out in the WHO Target Product Profile for a triage test.<sup>19</sup> 904 The box-and-whisker plots depict signature score (measured at enrolment) distribution by 905 symptom status (each dot represents a participant) in asymptomatic and symptomatic participants 906 with no TB, participants who progressed to incident TB, and participants with prevalent subclinical 907 (asymptomatic) and clinical (symptomatic) TB. Prevalent and incident TB comprised all primary 908 endpoint cases. Symptoms were recorded at the time of enrolment for participants without TB and 909 those with prevalent TB. p values for comparison of median signature scores between groups in 910 box-and-whisker plots were calculated with the Mann-Whitney U test and corrected for multiple 911 comparisons by use of the Benjamini-Hochberg Procedure.<sup>35</sup> Boxes depict the IQR, the midline 912 represents the median, and the whiskers indicate the IQR  $\pm$  (1.5 × IQR). 913 (S) Summary of signature diagnostic performance in order of primary endpoint AUC estimates.

- 914 The diagnostic AUC estimates in symptomatic and asymptomatic participant sub-groups are also
- 915 shown. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, and the black
- 916 dotted line indicates the lower bound of the 95% CI for the best performing signature for the
- 917 primary endpoint. The black dashed line (Figure 2S only) indicates an AUC cut-off of 0.5.
- 918

## Figures 3: Parsimonious signature prognostic performance for incident TB in peoplewithout HIV

- 921 Receiver operating characteristic curves depicting prognostic performance (area under the curve,
- AUC, with 95% CI) of the (A) Darboe11 (RISK11), (B) Francisco2, (C) Maertzdorf4 (DIAG4), (D)
- 923 Penn-Nicholson6 (RISK6), (E) Roe1 (BATF2), (F) Roe3, (G) Suliman4 (RISK4), (H) Sweeney3,
- and (I) Thompson5 (RESPONSE5) signatures for incident TB diagnosed on two or more liquid
- 925 culture-positive or Xpert MTB/RIF-positive sputum samples (primary endpoint) through 6, 12, and
- 926 15 months follow-up in the CORTIS-01 study of people without HIV.
- 927 The shaded areas represent the 95% Cls. The solid box depicts the optimal criteria (90%
- 928 sensitivity and 90% specificity) and the dashed box depicts the minimal criteria (75% sensitivity
- 929 and 75% specificity) set out in the WHO Target Product Profile for an incipient TB test.<sup>20</sup>
- 930 (J) Summary of signature prognostic performance in order of primary endpoint AUC estimates
- 931 through 12 months follow-up. The prognostic AUC estimates through 6 and 15 months are also
- shown. The midline indicates the AUC estimate, the error bars indicate the 95% CIs, and the black
- 933 dashed line indicates an AUC cut-off of 0.5.

934

# Figures 4: Parsimonious signature prognostic performance for prevalent TB in people livingwith HIV

- 937 Receiver operating characteristic curves depicting prognostic performance (area under the curve,
- AUC, with 95% CI) of the (A) Darboe11 (RISK11), (B) Francisco2, (C) Maertzdorf4 (DIAG4), (D)
- Penn-Nicholson6 (RISK6), (E) Roe1 (BATF2), (F) Roe3, (G) Suliman4 (RISK4), (H) Sweeney3,
- 940 and (I) Thompson5 (RESPONSE5) signatures for incident TB diagnosed on two or more liquid
- 941 culture-positive or Xpert MTB/RIF-positive sputum samples (primary endpoint) through 15 months
- follow-up in the CORTIS-HR study of people living with HIV.
- 943 The shaded areas represent the 95% Cls. The solid box depicts the optimal criteria (90%
- sensitivity and 90% specificity) and the dashed box depicts the minimal criteria (75% sensitivity
- <sup>945</sup> and 75% specificity) set out in the WHO Target Product Profile for an incipient TB test.<sup>20</sup>
- 946 (J) Summary of signature prognostic performance in order of primary endpoint AUC estimates
- 947 through 15 months follow-up. The midline indicates the AUC estimate, the error bars indicate the
- 948 95% Cls, the red dotted line indicates the lower bound of the 95% Cl for the best performing
- signature for the primary endpoint, and the black dashed line indicates an AUC cut-off of 0.5.
- 950

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

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

957 respiratory pathobionts sub-study (11 participants with secondary endpoint prevalent TB

- 958 excluded).
- 959 The box-and-whisker plots depict signature score distributions in participants with no upper
- 960 respiratory pathobionts (n=563), bacterial upper respiratory pathobionts only (n=352), viral upper
- 961 respiratory pathobionts only(n=41), both viral and bacterial upper respiratory pathobionts (n=33). p
- 962 values for comparison of median signature scores between groups in box-and-whisker plots were
- 963 calculated with the Mann-Whitney *U* test and corrected for multiple comparisons by use of the
- 964 Benjamini-Hochberg Procedure.<sup>35</sup> Only p-values below 0.1 are shown. Each dot represents a
- participant. Boxes depict the IQR, the midline represents the median, and the whiskers indicate the IQR  $\pm$  (1.5  $\times$  IQR).
- 967 The ROC curves depict performance (area under the curve, AUC, with 95% CI) of the
- 968 parsimonious signatures in differentiating between participants with viral upper respiratory
- 969 pathobionts and participants with bacterial upper respiratory pathobionts, between participants with
- 970 viral pathobionts and those with no pathobionts, or between participants with bacterial pathobionts
- and those with no pathobionts. The shaded areas represent the 95% Cls.







Figure 4



