



## Early View

### Original research article

## **Baseline IL-6 is a biomarker for unfavorable tuberculosis treatment outcomes: a multi-site discovery and validation study**

Akshay N. Gupte, Pavan Kumar, Mariana Araújo-Pereira, Vandana Kulkarni, Mandar Paradkar, Neeta Pradhan, Pradeep Menon, Padmapriya Darasini Chandrasekaran, Luke-Elizabeth Hanna, Shri Vijay Bala Yogendra Shivakumar, Neesha Rockwood, Elsa Du Bruyn, Rajesh Karyakarte, Sanjay Gaikwad, Robert Bollinger, Jonathan Golub, Nikhil Gupte, Vijay Viswanathan, Robert J. Wilkinson, Vidya Mave, Subash Babu, Hardy Kornfeld, Bruno B. Andrade, Amita Gupta

Please cite this article as: Gupte AN, Kumar P, Araújo-Pereira M, *et al.* Baseline IL-6 is a biomarker for unfavorable tuberculosis treatment outcomes: a multi-site discovery and validation study. *Eur Respir J* 2021; in press (<https://doi.org/10.1183/13993003.00905-2021>).

This manuscript has recently been accepted for publication in the *European Respiratory Journal*. It is published here in its accepted form prior to copyediting and typesetting by our production team. After these production processes are complete and the authors have approved the resulting proofs, the article will move to the latest issue of the ERJ online.

**Baseline IL-6 is a biomarker for unfavorable tuberculosis treatment outcomes: a multi-site discovery and validation study**

Akshay N. Gupte<sup>1,2</sup>, Pavan Kumar<sup>3</sup>, Mariana Araújo-Pereira<sup>4,5,6</sup>, Vandana Kulkarni<sup>2,7,8</sup>, Mandar Paradkar<sup>2,7,8</sup>, Neeta Pradhan<sup>2,7,8</sup>, Pradeep Menon<sup>3</sup>, Padmapriya Darasini<sup>3</sup>, Luke-Elizabeth Hanna<sup>3</sup>, Shri Vijay Bala Yogendra Shivakumar<sup>8</sup>, Neesha Rockwood<sup>9,10,11</sup>, Elsa Du Bruyn<sup>9,12</sup>, Rajesh Karyakarte<sup>13</sup>, Sanjay Gaikwad<sup>14</sup>, Robert Bollinger<sup>1,2</sup>, Jonathan Golub<sup>1,15</sup>, Nikhil Gupte<sup>1,2</sup>, Vijay Viswanathan<sup>16</sup>, Robert J. Wilkinson<sup>9,11,12,17</sup>, Vidya Mave<sup>1,2</sup>, Subash Babu<sup>18</sup>, Hardy Kornfeld<sup>19</sup>, Bruno B. Andrade<sup>4,5,6</sup>, Amita Gupta<sup>1,2</sup>

<sup>1</sup> Division of Infectious Diseases, Johns Hopkins University School of Medicine, Baltimore, USA

<sup>2</sup> Center for Clinical Global Health Education, Johns Hopkins University School of Medicine, Baltimore, USA

<sup>3</sup> National Institute for Research in Tuberculosis, Chennai, India

<sup>4</sup> Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (FIOCRUZ), Salvador, Brazil

<sup>5</sup> Multinational Organization Network Sponsoring Translational and Epidemiological Research, Salvador, Brazil

<sup>6</sup> Faculdade de Medicina, Universidade Federal da Bahia, Salvador, Brazil

<sup>7</sup> Byramjee-Jeejeebhoy Government Medical College-Johns Hopkins University Clinical Research Site, Pune, India

<sup>8</sup> Johns Hopkins India Private Limited, Pune, India

<sup>9</sup> Wellcome Centre for Infectious Disease Research in Africa, Institute of Infectious Disease and Molecular Medicine, University of Cape Town, Observatory, South Africa

<sup>10</sup> Department of Microbiology, Faculty of Medicine, University of Colombo, Colombo 8, Sri Lanka

<sup>11</sup> Department of Infectious Diseases, Imperial College London, W12 ONN, United Kingdom

<sup>12</sup> Department of Infectious Diseases, University of Cape Town, Observatory, South Africa

<sup>13</sup> Department of Microbiology, Byramjee-Jeejeebhoy Government Medical College, Pune, India

<sup>14</sup> Department of Pulmonary Medicine, Byramjee-Jeejeebhoy Government Medical College, Pune, India

<sup>15</sup> Center for Tuberculosis Research, Johns Hopkins University, Baltimore, USA

<sup>16</sup> Professor M. Viswanathan Diabetes Research Center, Chennai, India

<sup>17</sup> The Francis Crick Institute, 1 Midland Rd, London NW1 1AT

<sup>18</sup> National Institutes of Health – National Institute for Research in Tuberculosis – International Center for Excellence in Research, Chennai, India

<sup>19</sup> Division of Pulmonary Medicine, University of Massachusetts Medical School, Worcester, USA

**Corresponding author:** Akshay N. Gupte, agupte1@jhmi.edu

## **Abstract**

**Background:** Biomarkers of unfavorable tuberculosis treatment outcomes are needed to accelerate new drug and regimen development. Whether plasma cytokine levels can predict unfavorable tuberculosis treatment outcomes is unclear.

**Methods:** We identified and internally validated the association between 20 a-priori selected plasma inflammatory markers and unfavorable treatment outcomes of failure, recurrence and all-cause mortality among adults with drug-sensitive pulmonary tuberculosis in India. We externally validated these findings in two independent cohorts of predominantly diabetic and HIV coinfecting tuberculosis patients in India and South Africa, respectively.

**Results:** Pre-treatment IFN- $\gamma$ , IL-13 and IL-6 were associated with treatment failure in the discovery analysis. Internal validation confirmed higher pre-treatment IL-6 concentrations among failure cases compared to controls. External validation among predominantly diabetic tuberculosis patients found an association between pre-treatment IL-6 concentrations and subsequent recurrence and death. Similarly, external validation among predominantly HIV coinfecting tuberculosis patients found an association between pre-treatment IL-6 concentrations and subsequent treatment failure and death. In a pooled analysis of 363 tuberculosis cases from the Indian and South African validation cohorts, high pre-treatment IL-6 concentrations were associated with higher risk of failure (adjusted odds ratio [aOR]=2.16, 95%CI 1.08-4.33,  $p=0.02$ ), recurrence (aOR=5.36, 95%CI 2.48-11.57,  $p<0.001$ ) and death (aOR=4.62, 95%CI 1.95-10.95,  $p<0.001$ ). Adding baseline IL-6 to a risk-prediction model comprising

of low BMI, high smear grade and cavitation improved model performance by 15 percent (C-statistic of 0.66 versus 0.76,  $p=0.02$ ).

**Conclusion:** Pre-treatment IL-6 is a biomarker for unfavorable tuberculosis treatment outcomes. Future studies should identify optimal IL-6 concentrations for point-of-care risk prediction.

**Take home message:** Pre-treatment IL-6 is a biomarker for unfavorable tuberculosis treatment outcomes independent of disease severity and, improves the performance of risk-prediction models comprising of established clinical predictors.

## **Introduction**

Tuberculosis is the leading infectious cause of death worldwide with over 10 million new cases and 1.5 million deaths annually(1). Drug-sensitive cases require six months of multi-drug therapy for durable cure. Although effective in the majority of cases, implementing this relatively long and complex treatment regimen can increase risk of default and overutilization of health system resources. Developing shorter and simpler tuberculosis treatment regimens is therefore a research priority. Biomarkers have the potential to accelerate new drug and regimen discovery by identifying tuberculosis patients at high risk of unfavorable treatment outcomes who can preferentially be enrolled into early phase clinical trials, or conversely, characterize a subgroup of low-risk patients who may benefit from shorter duration of therapy(2).

Sputum culture conversion at two months of treatment is the most widely used biomarker of treatment response; however, its validity in predicting subsequent clinical outcomes has been questioned in recent treatment shortening trials(3, 4). Blood based biomarkers have an advantage as they overcome the difficulty of reliably obtaining high quality sputum and can be translated into point-of-care tests. A recent study by Kumar et al identified an association between a combination of six chemokines and unfavorable tuberculosis treatment outcomes, suggesting a role of plasma inflammatory markers as predictors of unfavorable treatment outcomes(5). However, translating such a “signature” comprising of multiple and often correlated immune markers, each with their different concentration threshold for optimal risk-prediction, to a point-of-care test is challenging. Whether a parsimonious selection of inflammatory markers can predict unfavorable treatment outcomes in tuberculosis patients is unclear.

We have previously identified plasma cytokines associated with lung injury in pulmonary tuberculosis cases in India(6). We now report on a multi-site discovery-validation analysis in India and South Africa to identify plasma cytokines associated with treatment failure, recurrence and death.

## **Methods**

### **Study design and population**

We conducted a prospective cohort study to discover plasma cytokines associated with unfavorable treatment outcomes among adult drug-sensitive pulmonary tuberculosis patients enrolled in the Cohort for Tuberculosis Research by the Indo-US Medical Partnership (CTRIUMPH) study in Pune, India(7). We internally validated these results among CTRIUMPH participants not previously included in the discovery cohort by nesting a 1:1 age- and sex-matched case-control analysis among adult drug-sensitive pulmonary tuberculosis patients who failed treatment (cases) and those who were cured (controls). We subsequently conducted external validation analyses in two independent cohorts from India and South Africa. The Indian external validation cohort comprised of predominantly diabetic drug-sensitive pulmonary tuberculosis patients from the Effect of Diabetes on Tuberculosis Severity (EDOTS) study in Chennai, India(8). In this cohort, we nested a 1:2 age-, sex- and body-mass-index (BMI)-matched case-control analysis among adult drug-sensitive pulmonary tuberculosis patients who experienced a composite unfavorable treatment outcome (cases) of failure, recurrence or death, and those with recurrence free cure over 18 months of follow-up (controls). The South African external validation cohort comprised of predominantly HIV coinfecting drug-

sensitive pulmonary tuberculosis patients from Khayelitsha, South Africa(9, 10). In this cohort, we conducted an unmatched case-control analysis among adult drug-sensitive pulmonary tuberculosis patients who experienced a composite unfavorable treatment outcome (cases) of failure, recurrence or death, and those with recurrence free cure over 18 months of follow-up (controls). Finally, we pooled data from the Indian and South African cohorts for a combined validation analysis. Detailed descriptions of the discovery, internal validation, Indian external validation and South African external validation cohorts are provided in Supplementary Appendix 1. In addition to different study populations, we used different laboratories for measuring cytokine concentrations for the internal and external validation studies to assess generalizability of our findings (Supplementary Table-1). This study was approved by the Ethics Committees at the participating study sites in India and South Africa.

### **Cytokine measurement**

For the discovery analysis, plasma samples collected at treatment initiation, 2 months and 6 months underwent cytokine testing, in duplicates, using multiplex ELISA on Luminex assay (Bio-Rad, USA) at the National Institutes of Health (NIH) – National Institute for Research in Tuberculosis (NIRT) – International Center for Excellence in Research (ICER) laboratory in Chennai, India. Cytokines analyzed were selected a-priori for their role in the host inflammatory response to *Mycobacterium tuberculosis* (interferon gamma [INF- $\gamma$ ], tumor necrosis factor alpha [TNF- $\alpha$ ], interleukin [IL]-1 $\beta$ , IL-4, IL-6, IL-10, IP-10, IL-12, IL-13 and IL-17)(11), lung pathology (matrix metalloproteinases [MMP]-1, MMP-3, MMP-7, tissue inhibitor of metalloprotease [TIMP]-1, TIMP-2, TIMP-3,



TIMP-4)(12), and fibrous remodeling (transforming growth factor beta [TGF $\beta$ ]-1, TGF $\beta$ -2, TGF $\beta$ -3)(13). IL-6, IL-13 and IFN- $\gamma$  measured at treatment initiation were statistically significantly associated with treatment failure in the discovery analysis and were therefore selected for internal validation. For the internal validation analysis, plasma samples collected at treatment initiation underwent cytokine testing, in duplicates, using multiplex ELISA by Luminex assay (Bio-Rad, USA) at the Byramjee-Jeejeebhoy Government Medical College (BJGMC) laboratory in Pune, India. IL-6 measured at treatment initiation was statistically significantly associated with treatment failure during internal validation and was therefore selected for independent external validation. For the Indian external validation cohort, plasma samples collected at treatment initiation underwent IL-6 testing by Luminex assay (R&D Systems, USA) at the NIH-NIRT-ICER laboratory in Chennai, India. For the South African external validation cohort, plasma samples collected at tuberculosis treatment initiation underwent IL-6 testing by Luminex assay (MilliporeSigma, USA) at the Wellcome Centre for Infectious Disease Research, South Africa. The lower limit of IL-6 detection was 0.31 pg/mL and a nonlinear curve fit model was used to estimate IL-6 values less than the lower standard values.

## **Outcomes**

Unfavorable tuberculosis treatment outcomes included failure, recurrence and death. Treatment failure was defined as culture confirmation of *M tuberculosis* during the last two months of treatment. Recurrence was defined in a tuberculosis patient who did not fail treatment but was subsequently found to have culture positive tuberculosis during 18 months of post-treatment follow-up. Death included all-cause mortality.

## **Statistical analysis**

**Discovery:** We compared cytokine expression at treatment initiation, 2 months and 6 months between tuberculosis cases with and without an unfavorable treatment outcome using one-way hierarchical cluster analysis by Ward's method with 100x bootstrap. Median concentrations of cytokines, stratified by treatment outcomes and duration, were used to describe the overall change in cytokine expression in response to tuberculosis treatment. Cytokine concentrations were  $\log_{10}$ -transformed, or z-score normalized for analysis. Data were compared using the Mann-Whitney U test and P-values were adjusted for multiple comparisons using the Holm-Bonferroni method(15).

**Internal and external validation:** We compared  $\log_{10}$ -transformed cytokine concentrations between participants with and without an unfavorable treatment outcome using the Wilcoxon sign-rank test. Receiver Operator Characteristic Curve (ROC) analysis was used to assess the ability of individual cytokines to discriminate between participants with and without unfavorable treatment outcomes. Random-effects linear regression, with matched case-control pairs as random effects, was used to compare cytokine concentrations between participants with and without unfavorable treatment outcomes. Multivariable analyses were adjusted for tuberculosis disease severity assessed by BMI, chest radiography (CXR) score including cavitation, and sputum smear grade. We further adjusted for pre-treatment illness duration to account for differences in the timing of patient presentation relative to disease onset and its possible impact on cytokine concentrations, glycated hemoglobin (HbA1c) among participants

with diabetes and, CD4 cell counts and receipt of antiretroviral therapy (ART) among HIV coinfecting participants.

**Pooled validation:** We pooled data from the internal and external validation cohorts to measure the association between baseline IL-6 concentrations and unfavorable tuberculosis treatment outcomes. Since each of the three validation cohorts utilized different laboratories, ELISA kits and testing protocols, there was substantial variability in picogram/milliliter (pg/mL) IL-6 quantification across the study sites. To account for this variability, we used z-scores and percentiles to standardize IL-6 concentrations across the three validation cohorts. For each validation cohort, we calculated a z-score for pg/mL IL-6 concentrations using the formula  $z = (\text{observed concentration} - \text{mean concentration}) / \text{standard deviation}$ . This standardized z-score was used for ROC analysis in the pooled validation cohort. Similarly, we classified “high” IL-6 concentrations as having pg/mL concentrations in the highest quartile for a given study. “Low” IL-6 concentrations were defined as having pg/mL concentrations in the lower three quartiles for a given study. To account for clustering by study location and case-control matching, we used random-effects logistic regression with bootstrap confidence intervals, to measure the association between “high” baseline IL-6 concentrations and unfavorable tuberculosis treatment outcomes of failure, recurrence or death. Multivariable analyses adjusted for age, sex, BMI, smear grade, CXR score including cavitation, pre-treatment illness duration, HIV and diabetes. Furthermore, cavitory disease, low BMI and higher smear grade have previously been shown to predict unfavorable tuberculosis treatment outcomes(16). Therefore, we calculated the

incremental gain in discriminatory ability, measured by the C-statistic, by adding baseline IL-6 to a risk-prediction model comprising of cavitation, BMI and smear grade.

## **Results**

Baseline characteristics of the discovery sub-cohort did not differ significantly from the full cohort and are summarized in Supplementary Table-2. Of the 30 participants selected in the discovery sub-cohort, 4 (13%) failed treatment, and none had recurrence or died. Participants who failed treatment had significantly higher baseline concentrations of IL-6, INF- $\gamma$  and IL-13 relative to those without failure ( $p < 0.001$  for all); however, we did not find statistically significant differences in cytokine levels measured at 2 and 6 months of treatment (Figure-1). The internal validation cohort comprised of 20 participants with culture confirmed treatment failure and 20 age- and sex-matched participants with culture confirmed cure. Overall, 3 (8%) participants were HIV coinfecting and 3 (8%) had diabetes. Except for lower median BMI (16.0 vs 17.4 kg/m<sup>2</sup>,  $p = 0.02$ ) and longer pre-treatment illness duration (35 vs 30 days,  $p = 0.03$ ), cases were comparable to controls with respect to their baseline characteristics (Supplementary Table-3). After adjusting for markers of disease severity and pre-treatment illness duration, cases had significantly higher IL-6 concentrations at baseline relative to controls (0.26 log-higher concentration, 95%CI 0.01 to 0.52,  $p = 0.04$ ) with an area under the curve (AUC) of 0.70 (95%CI 0.52-0.88). We did not find differences in INF- $\gamma$  and IL-13 concentrations by treatment outcomes (Table-1 and Figure-2).

The Indian external validation cohort was comprised of 72 cases with an unfavorable treatment outcome and 122 age-, sex- and BMI-matched controls with recurrence free cure. Overall, 115 (59%) participants had diabetes and HIV was an exclusion criterion. Cases were comparable to controls with respect to their baseline characteristics, except for a higher proportion of ever-smokers (Supplementary Table-4). Overall, 18 (9%) participants failed treatment, 35 (18%) had recurrence, and 19 (10%) died. Cases with an unfavorable treatment outcome had significantly higher concentrations of IL-6 at baseline relative to controls (0.26 log-higher concentration, 95%CI 0.17 to 0.36,  $p < 0.001$ ), after adjusting for markers of disease severity, pre-treatment illness duration, smoking status and diabetes. Results were similar when restricted to tuberculosis patients with diabetes and adjusting for HbA1c levels; cases experiencing recurrence or death had higher baseline IL-6 concentrations compared to controls (Table-2 and Figure-3). IL-6 could effectively discriminate tuberculosis patients with and without unfavorable treatment outcomes with an AUC of 0.73 (95%CI 0.66-0.80) for all participants and 0.73 (95%CI 0.64-0.83) when restricted to participants with diabetes (Figure-3).

Of the 129 participants enrolled in the South African external validation cohort, 76 (59%) had HIV, 29 (38%) were receiving antiretroviral therapy (ART) and the median (IQR) CD4 count was 192 (66-366) cells/mm<sup>3</sup> (Supplementary Table-5). Overall, 18 (14%) participants experienced an unfavorable treatment outcome; 9 (7%) failed treatment, 4 (3%) had recurrence and 5 (4%) died. After adjusting for age, sex, HIV status, ART and markers of disease severity, participants who experienced an

unfavorable treatment outcome had significantly higher concentrations of IL-6 at baseline compared to those with recurrence free cure (0.38 log-higher concentration, 95%CI 0.13 to 0.63,  $p=0.003$ ). However, we found marginally significant associations between higher baseline IL-6 concentrations and unfavorable tuberculosis treatment outcomes when restricted to HIV coinfecting participants (Table-3). IL-6 could discriminate tuberculosis patients with and without unfavorable treatment outcomes with an AUC of 0.69 (95%CI 0.55-0.85) for all participants and 0.66 (95%CI 0.42-0.85) when restricted to HIV coinfecting participants (Figure-4).

The pooled validation from Indian and South African cohorts comprised of 363 participants experiencing 110 unfavorable outcomes: 47 (13%) had failure, 39 (11%) had recurrence and 24 (6%) died. High IL-6 concentration at baseline was present in 91 (25%) participants and was associated with 3-fold higher odds of any subsequent unfavorable treatment outcome (95%CI 2.03-5.89,  $p<0.001$ ) after adjusting for age, sex, BMI, CXR score including cavitation, smear grade, pre-treatment illness duration, HIV, diabetes and correlations within each contributing study (Table-4). Similarly, high baseline IL-6 concentrations were significantly associated with subsequent failure (adjusted odds ratio [aOR]=2.22, 95%CI 1.04-4.75,  $p=0.03$ ), recurrence (aOR=5.92, 95%CI 2.52-13.90,  $p<0.001$ ) and death (aOR=5.53, 95%CI 2.08-14.68,  $p<0.001$ ) in adjusted analyses.

Low BMI, high smear grade ( $\geq 2+$ ) and cavitation on CXR were poor baseline predictors of unfavorable tuberculosis treatment outcomes in our cohort. A risk-prediction model comprising of these variables had a C-statistic (AUC) of 0.66 (95%CI 0.56-0.77) for discriminating between tuberculosis patients with and without unfavorable

treatment outcomes. However, baseline IL-6 could discriminate tuberculosis patients with and without unfavorable treatment outcomes with an AUC of 0.71 (95%CI 0.65-0.77); a z-score cut-off of -0.32 identified by the Youden's method had a sensitivity of 76%, specificity of 60%, and a negative predictive value of 85% for ruling out subsequent unfavorable tuberculosis treatment outcomes. Furthermore, the addition of baseline IL-6 improved performance of the prediction model comprising of low BMI, high smear grade and cavitation by 15 percent (C-statistic=0.76, 95%CI 0.67-0.85, p=0.02) (Supplementary Table-6).

## **Discussion**

Our discovery-validation analyses among geographically and epidemiologically diverse populations of pulmonary tuberculosis cases found an association between higher baseline IL-6 concentrations in plasma and subsequent treatment failure, recurrence and death. This association was generalizable despite different testing laboratories, comorbidities of diabetes and HIV, and after adjusting for disease severity. Furthermore, the inclusion of baseline IL-6 significantly improved the performance of established risk-prediction models for poor clinical outcomes. Our data support the role of IL-6 as a biomarker for unfavorable tuberculosis treatment outcomes.

IL-6 exhibits a pleiotropic immunoregulatory role by promoting neutrophil and macrophage recruitment and survival, inducing the acute phase response in the liver, and facilitating tissue damage by stimulating protease secretion and matrix deposition(17, 18). IL-6 has long been considered a non-specific marker of tuberculosis disease activity and, prior studies have identified a correlation between IL-6

concentrations, bacterial burden and lung pathology(6, 19, 20). However, few studies have reported an association between IL-6 and tuberculosis treatment outcomes. A secondary analysis of an early phase treatment trial reported correlations between tuberculosis disease severity and baseline IL-6 concentrations, and between greater declines in IL-6 concentrations and sputum culture conversion during the first two months of treatment(21). However, this study did not report on final treatment outcomes. Across the three independent cohorts analyzed in our study, we found a consistent association between higher baseline IL-6 concentrations and subsequent unfavorable tuberculosis treatment outcomes. This association remained significant after adjusting for markers of disease severity and suggests an independent predictive role of baseline IL-6 for risk-stratification in tuberculosis patients.

While our study found an association of baseline IL-6 concentrations with recurrent tuberculosis in participants with and without diabetes, we did not find a similar association among HIV coinfecting participants. This finding is inconsistent with a recent study by Sivro et al., which reported an association between IL-6 concentrations, measured after completion of tuberculosis treatment, and recurrent disease among ART naïve HIV coinfecting patients(22). This inconsistency may be due to differences in the underlying study populations and timing of IL-6 measurement – treatment initiation in our study compared to after treatment completion in the Sivro et al study. However, our sample size for HIV-restricted analysis was limited and further validation in larger cohorts of HIV coinfecting tuberculosis patients is needed. Conversely, we found higher baseline IL-6 concentrations among tuberculosis patients with diabetes who had recurrence or died; we did not find similar differences by treatment failure. The precise



reasons for this are unclear. The internal validation cohort was specifically designed to examine an outcome of treatment failure, while the external validation cohorts did not perform a-priori sampling on treatment outcomes. A relatively fewer number of treatment failure cases in the external validation cohorts may explain a weaker association of baseline IL-6 with treatment failure. Furthermore, diabetes is characterized by hyperinflammation and higher overall IL-6 concentrations among participants with diabetes at the start of tuberculosis treatment could explain this finding(23-25). Nevertheless, further study is needed explore longitudinal changes in the immunological profile of tuberculosis patients with diabetes who fail treatment.

A consistent finding in our study, regardless of HIV or diabetes comorbidity, was the association between baseline IL-6 with all-cause mortality. The ability of IL-6 to predict mortality has previously been reported in several non-communicable and infectious diseases, and more recently in COVID-19(26-30). While studies among tuberculosis patients with advanced HIV do suggest a role of IL-6 in explaining excess mortality(31, 32), similar associations among HIV uninfected tuberculosis patients have yet to be investigated. Our study addresses this knowledge gap by reporting a consistent and significant association between high baseline IL-6 concentrations and subsequent all-cause mortality in HIV uninfected tuberculosis patients with and without diabetes. Importantly, majority of deaths in our study occurred during tuberculosis treatment and, the association between IL-6 and mortality was independent of baseline disease severity and duration of symptomatic illness prior to testing. These data suggest an independent role of IL-6 mediated immune pathways in mortality during tuberculosis treatment.

An important limitation of our study was the relatively small discovery cohort and a limited selection of cytokines. We selected cytokines a-priori based on their biological plausibility; however, a non-targeted approach may identify additional cytokines predictive of unfavorable outcomes. Further, our randomly selected discovery cohort was exploratory in nature and did not include participants with recurrence or death. Therefore, we are likely underpowered to detect a statistically significant association of a smaller magnitude for selected cytokines. We did not objectively assess treatment adherence. However, we did measure self-reported treatment adherence and all participants reported taking over 95% of their prescribed doses. Finally, each of the validation cohorts used a different laboratory, ELISA kits and testing protocols which led to a high variability in IL-6 quantification. Therefore, we were unable to identify an IL-6 pg/mL concentration cut-off for point-of-care risk-prediction.

Despite these limitations, our study found a significant and generalizable association between high baseline IL-6 concentrations and subsequent unfavorable tuberculosis treatment outcomes in a relatively large and epidemiologically diverse population. Although the overall discriminatory ability of IL-6 in our study was modest, it was nevertheless consistent across the Indian and South African validation cohorts. Identifying concentration thresholds to optimize either sensitivity or specificity could facilitate the use of baseline IL-6 as a rule-out or rule-in test, respectively, for risk-stratification at treatment initiation. Such an approach may be important given the poor performance of low BMI, high smear grade and cavitory disease, clinical characteristics that have recently been shown to identify hard-to-treat phenotypes of tuberculosis patients in clinical trial settings(16), in predicting unfavorable treatment outcomes in

programmatic settings as seen in this study. For instance, the inclusion of baseline IL-6 in a risk-prediction model comprising of these clinical variables significantly improved performance of the prediction model by nearly 15 percent in our study, supporting a strategy of combining baseline IL-6 with established clinical predictors of unfavorable outcomes for risk-stratification in tuberculosis. Furthermore, in contrast to prior studies which identified predictive signatures comprising of multiple immune markers, our data identified a single cytokine predictive of unfavorable tuberculosis treatment outcomes, potentially simplifying its translation to a point-of-care assay. Future studies should focus on standardizing IL-6 measurements across laboratories for point-of-care translation, and identify optimal predictive thresholds of IL-6 concentrations for risk-stratification.

**Funding:** Data in this manuscript were collected as part of the Regional Prospective Observational Research for Tuberculosis (RePORT) India Consortium. This project has been funded in whole or in part with Federal funds from the Government of India's (GOI) Department of Biotechnology (DBT), the Indian Council of Medical Research (ICMR), the United States National Institutes of Health (NIH), National Institute of Allergy and Infectious Diseases (NIAID), Office of AIDS Research (OAR), and distributed in part by CRDF Global. Research reported in this publication was also supported by NIAID of the NIH under award number R01AI097494 and UM1AI069465, the Wyncote Foundation, the Gilead Foundation and the Ujala Foundation, Newton Square PA, USA. This research was also funded in part by a 2019 developmental grant from the Johns Hopkins University Center for AIDS Research, an NIH funded program (1P30AI094189), which is supported by the following NIH Co-Funding and Participating Institutes and Centers: NIAID, NCI, NICHD, NHLBI, NIDA, NIA, NIGMS, NIDDK, NIMHD. ANG was supported by the National Institute Of Allergy And Infectious Diseases of the National Institutes of Health under Award Number K99AI151094. MAP received a fellowship from Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Finance code: 001). The work of BBA and RJW was supported by a grant from NIAID (U01-AI115940), by the Intramural Research Program of the Oswaldo Cruz Foundation (FIOCRUZ) and by the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil. RJW is supported by The Francis Crick Institute which receives support from MRC (FC001218), Cancer Research UK (FC001218) and Wellcome (FC001218). Additional support was provided by the Wellcome Trust (203135, 104803), the European and Developing Countries Clinical Trials Partnership

(SRIA 2015-1065) and the Foundation for the National Institutes of Health (WILK116PTB). The funders had no role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the article for publication. The contents of this publication are solely the responsibility of the authors and do not represent the official views of the DBT, ICMR, MRC, NIH, or CRDF Global. Any mention of trade names, commercial projects, or organizations does not imply endorsement by any of the sponsoring organizations. The authors also acknowledge support from Persistent Systems in kind.

## References

1. Global Tuberculosis Report Geneva: World Health Organization 2020.
2. Wallis RS, Kim P, Cole S, Hanna D, Andrade BB, Maeurer M, et al. Tuberculosis biomarkers discovery: developments, needs, and challenges. *The Lancet Infectious diseases*. 2013;13(4):362-72.
3. Phillips PP, Fielding K, Nunn AJ. An evaluation of culture results during treatment for tuberculosis as surrogate endpoints for treatment failure and relapse. *PloS one*. 2013;8(5):e63840.
4. Wallis RS, Peppard T, Hermann D. Month 2 culture status and treatment duration as predictors of recurrence in pulmonary tuberculosis: model validation and update. *PloS one*. 2015;10(4):e0125403.
5. Kumar NP, Moideen K, Nancy A, Viswanathan V, Thiruvengadam K, Nair D, et al. Plasma chemokines are baseline predictors of unfavorable treatment outcomes in pulmonary tuberculosis. *Clin Infect Dis*. 2020.
6. Gupte AN, Selvaraju S, Gaikwad S, et al. Higher IL-6 Levels and Changes in TGF- $\beta$  are Associated with Lung Impairment in Pulmonary Tuberculosis. *ERJ Open Res* 2020; in press (<https://doi.org/10.1183/23120541.00390-2020>). .
7. Gupte A, Padmapriyadarsini C, Mave V, Kadam D, Suryavanshi N, Shivakumar SV, et al. Cohort for Tuberculosis Research by the Indo-US Medical Partnership (CTRIUMPH): protocol for a multicentric prospective observational study. *BMJ open*. 2016;6(2):e010542.
8. Kornfeld H, West K, Kane K, Kumpatla S, Zacharias RR, Martinez-Balzano C, et al. High Prevalence and Heterogeneity of Diabetes in Patients With TB in South India: A

Report from the Effects of Diabetes on Tuberculosis Severity (EDOTS) Study. *Chest*. 2016;149(6):1501-8.

9. Rockwood N, Costa DL, Amaral EP, Du Bruyn E, Kubler A, Gil-Santana L, et al. Mycobacterium tuberculosis Induction of Heme Oxygenase-1 Expression Is Dependent on Oxidative Stress and Reflects Treatment Outcomes. *Front Immunol*. 2017;8:542.

10. Rockwood N, Sirgel F, Streicher E, Warren R, Meintjes G, Wilkinson RJ. Low Frequency of Acquired Isoniazid and Rifampicin Resistance in Rifampicin-Susceptible Pulmonary Tuberculosis in a Setting of High HIV-1 Infection and Tuberculosis Coprevalence. *The Journal of infectious diseases*. 2017;216(6):632-40.

11. O'Garra A, Redford PS, McNab FW, Bloom CI, Wilkinson RJ, Berry MP. The immune response in tuberculosis. *Annu Rev Immunol*. 2013;31:475-527.

12. Ong CW, Elkington PT, Friedland JS. Tuberculosis, pulmonary cavitation, and matrix metalloproteinases. *American journal of respiratory and critical care medicine*. 2014;190(1):9-18.

13. Tatler AL, Jenkins G. TGF-beta activation and lung fibrosis. *Proc Am Thorac Soc*. 2012;9(3):130-6.

14. J. AITCHISON, Principal component analysis of compositional data, *Biometrika*, Volume 70, Issue 1, April 1983, Pages 57–65, <https://doi.org/10.1093/biomet/70.1.57>.

15. Holm S. A Simple Sequentially Rejective Multiple Test Procedure. *Scandinavian Journal of Statistics*. 1979;6(2):65-70.

16. Imperial MZ, Nahid P, Phillips PPJ, Davies GR, Fielding K, Hanna D, et al. A patient-level pooled analysis of treatment-shortening regimens for drug-susceptible pulmonary tuberculosis. *Nat Med*. 2018;24(11):1708-15.

17. Rincon M. Interleukin-6: from an inflammatory marker to a target for inflammatory diseases. *Trends Immunol.* 2012;33(11):571-7.
18. Hunter CA, Jones SA. IL-6 as a keystone cytokine in health and disease. *Nat Immunol.* 2015;16(5):448-57.
19. Casarini M, Ameglio F, Alemanno L, Zangrilli P, Mattia P, Paone G, et al. Cytokine levels correlate with a radiologic score in active pulmonary tuberculosis. *American journal of respiratory and critical care medicine.* 1999;159(1):143-8.
20. Mesquita ED, Gil-Santana L, Ramalho D, Tonomura E, Silva EC, Oliveira MM, et al. Associations between systemic inflammation, mycobacterial loads in sputum and radiological improvement after treatment initiation in pulmonary TB patients from Brazil: a prospective cohort study. *BMC Infect Dis.* 2016;16:368.
21. Sigal GB, Segal MR, Mathew A, Jarlsberg L, Wang M, Barbero S, et al. Biomarkers of Tuberculosis Severity and Treatment Effect: A Directed Screen of 70 Host Markers in a Randomized Clinical Trial. *EBioMedicine.* 2017;25:112-21.
22. Sivro A, McKinnon LR, Yende-Zuma N, Gengiah S, Samsunder N, Abdool Karim SS, et al. Plasma Cytokine Predictors of Tuberculosis Recurrence in Antiretroviral-Treated Human Immunodeficiency Virus-infected Individuals from Durban, South Africa. *Clin Infect Dis.* 2017;65(5):819-26.
23. Restrepo BI, Fisher-Hoch SP, Pino PA, Salinas A, Rahbar MH, Mora F, et al. Tuberculosis in poorly controlled type 2 diabetes: altered cytokine expression in peripheral white blood cells. *Clin Infect Dis.* 2008;47(5):634-41.
24. Kumar NP, Sridhar R, Banurekha VV, Jawahar MS, Fay MP, Nutman TB, et al. Type 2 diabetes mellitus coincident with pulmonary tuberculosis is associated with



heightened systemic type 1, type 17, and other proinflammatory cytokines. *Ann Am Thorac Soc.* 2013;10(5):441-9.

25. Prada-Medina CA, Fukutani KF, Pavan Kumar N, Gil-Santana L, Babu S, Lichtenstein F, et al. Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. *Sci Rep.* 2017;7(1):1999.

26. Baune BT, Rothermundt M, Ladwig KH, Meisinger C, Berger K. Systemic inflammation (Interleukin 6) predicts all-cause mortality in men: results from a 9-year follow-up of the MEMO Study. *Age (Dordr).* 2011;33(2):209-17.

27. Lindmark E, Diderholm E, Wallentin L, Siegbahn A. Relationship between interleukin 6 and mortality in patients with unstable coronary artery disease: effects of an early invasive or noninvasive strategy. *JAMA.* 2001;286(17):2107-13.

28. Srisangthong P, Wongs A, Kittiworawitkul P, Wattanathum A. Early IL-6 response in sepsis is correlated with mortality and severity score. *Crit Care.* 2013;17(Suppl 2):P34. doi:10.1186/cc11972.

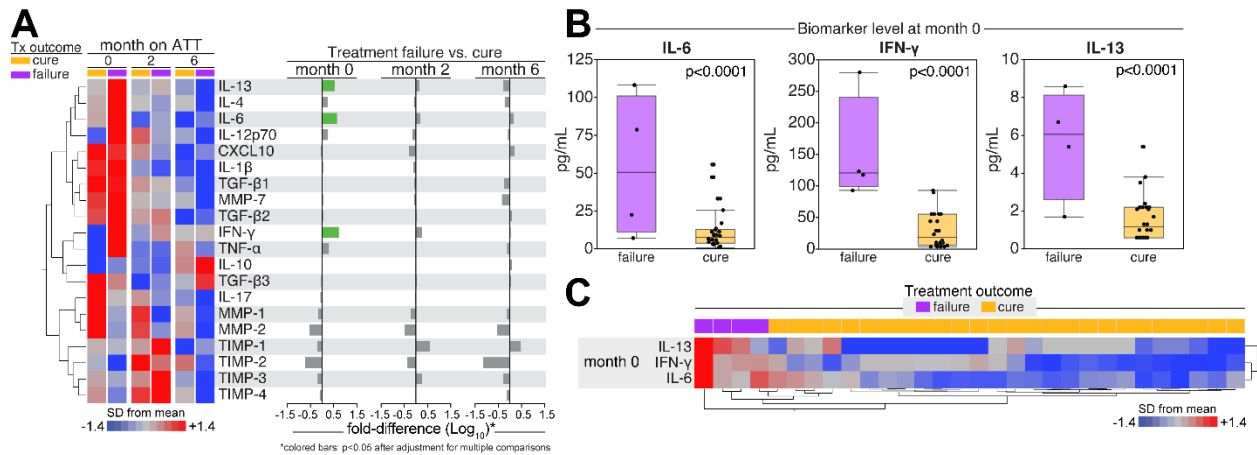
29. Leisman DE, Ronner L, Pinotti R, Taylor MD, Sinha P, Calfee CS, et al. Cytokine elevation in severe and critical COVID-19: a rapid systematic review, meta-analysis, and comparison with other inflammatory syndromes. *The Lancet Respiratory medicine.* 2020;8(12):1233-44.

30. McElvaney OJ, McEvoy NL, McElvaney OF, Carroll TP, Murphy MP, Dunlea DM, et al. Characterization of the Inflammatory Response to Severe COVID-19 Illness. *American journal of respiratory and critical care medicine.* 2020;202(6):812-21.

31. Manabe YC, Andrade BB, Gupte N, Leong S, Kintali M, Matoga M, et al. A Parsimonious Host Inflammatory Biomarker Signature Predicts Incident TB and Mortality in Advanced HIV. *Clin Infect Dis*. 2019.

32. Ravimohan S, Tamuhla N, Steenhoff AP, Letlhogile R, Nfanyana K, Bellamy SL, et al. Immunological profiling of tuberculosis-associated immune reconstitution inflammatory syndrome and non-immune reconstitution inflammatory syndrome death in HIV-infected adults with pulmonary tuberculosis starting antiretroviral therapy: a prospective observational cohort study. *The Lancet Infectious diseases*. 2015;15(4):p429–38.

**Figure 1.** Cytokines associated with treatment failure in the discovery cohort.

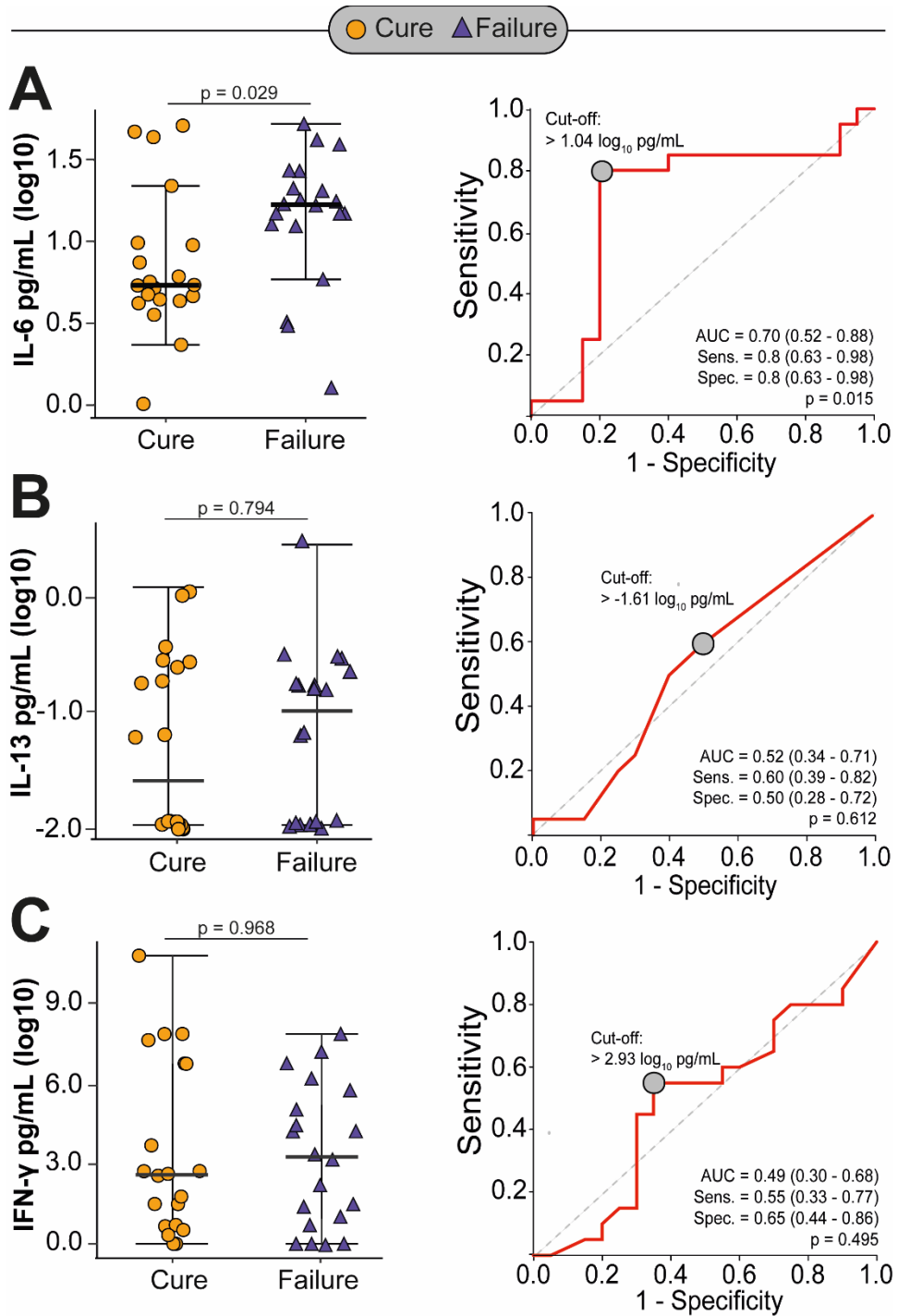


**Panel A:** Differences in cytokine concentrations (z-score standardized and  $\log_{10}$ -transformed) between participants who failed treatment and those who were cured, stratified by duration of treatment.

**Panel B:** Differences in absolute cytokine concentrations (pg/mL) at enrollment between participants who failed treatment and those who were cured.

**Panel C:** A two-way unsupervised hierarchical cluster analysis using concentrations of IL-6, IFN- $\gamma$  and IL-13 measured at enrollment to identify a unique combined profile of biomarker protein expression that could distinguish participants based on treatment outcome.

**Figure 2.** Difference in IL-6 (Panel A), IL-13 (Panel B) and IFN- $\gamma$  (Panel C) concentrations between tuberculosis treatment failures and cures, and associated area under the curve (AUC) for failure-cure classification in the internal validation cohort.

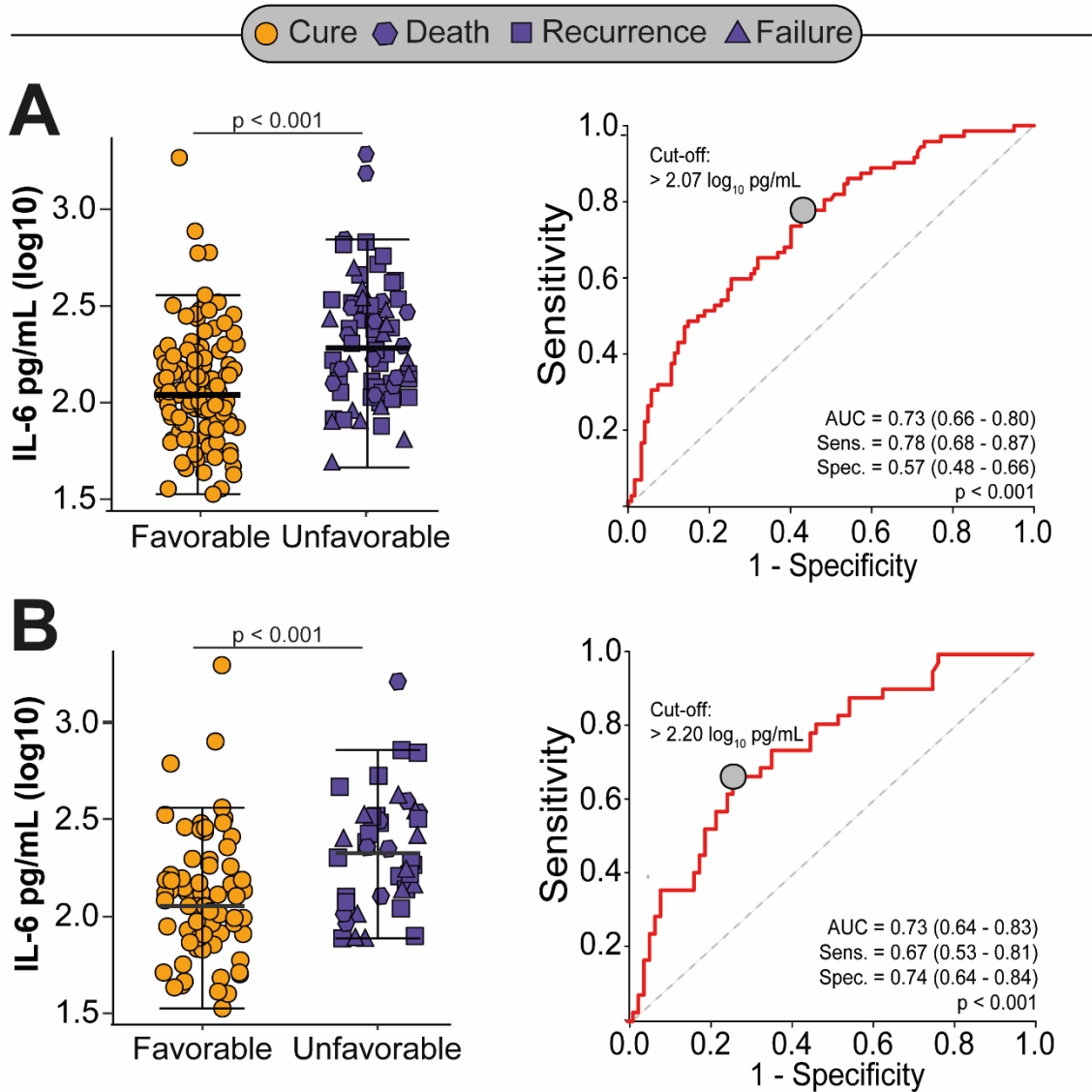


**Table 1.** Difference in cytokine concentrations between tuberculosis patients who failed treatment (cases) and those who were cured (controls) in the internal validation cohort.

<b>Cytokines</b>	<b>Log<sub>10</sub> difference in baseline cytokine levels comparing cases to controls</b>			
	<i>Unadjusted difference (95% CI)</i>	<i>p-value</i>	<i>Adjusted difference (95% CI)</i>	<i>p-value</i>
<b>IL-6</b>	0.29 (0.06 to 0.52)	0.01	0.26 (0.01 to 0.52)	0.04
<b>IFN-γ</b>	-0.12 (-0.68 to 0.42)	0.65	-0.38 (-0.98 to 0.21)	0.21
<b>IL-13</b>	0.09 (-0.20 to 0.39)	0.52	0.13 (-0.22 to 0.50)	0.45

Adjusted analysis account for BMI, CXR score including cavitation, smear grade and pre-treatment illness duration. Age and sex are adjusted by study design. CI – confidence interval.

**Figure 3.** Difference in IL-6 concentrations between tuberculosis patients with and without unfavorable treatment outcomes, and associated area under the curve (AUC) among all participants (Panel A) and restricted to patients with diabetes only (Panel B) in the Indian external validation cohort.

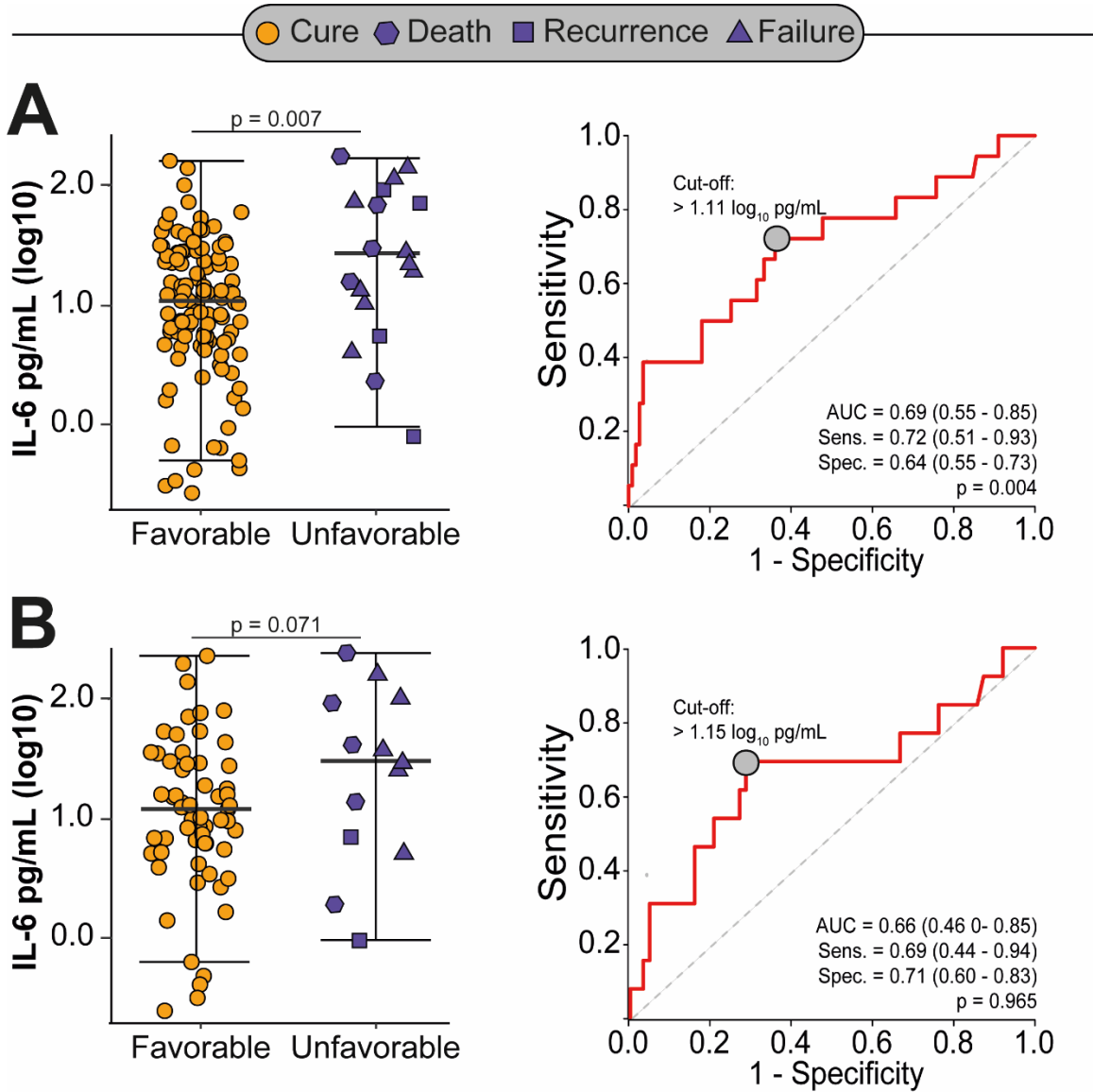


**Table 2.** Difference in IL-6 concentrations between with tuberculosis patients with and without unfavorable treatment outcomes in the Indian external validation cohort.

Treatment outcomes	Log <sub>10</sub> difference in baseline cytokine levels comparing cases to controls			
	<i>Unadjusted difference (95%CI)</i>	<i>p-value</i>	<i>Adjusted difference (95%CI)</i>	<i>p-value</i>
<b>All participants</b>				
Composite	0.24 (0.16 to 0.32)	<0.001	0.26 (0.17 to 0.36)	<0.001
Failure	0.10 (-0.02 to 0.24)	0.10	0.16 (-0.02 to 0.36)	0.07
Recurrence	0.26 (0.17 to 0.35)	<0.001	0.27 (0.16 to 0.38)	<0.001
Death	0.32 (0.18 to 0.46)	<0.001	0.32 (0.14 to 0.50)	<0.001
<b>Diabetes only</b>				
Composite	0.23 (0.12 to 0.35)	<0.001	0.22 (0.07 to 0.37)	0.01
Failure	0.15 (-0.03 to 0.34)	0.11	0.15 (-0.12 to 0.43)	0.27
Recurrence	0.22 (0.10 to 0.36)	0.001	0.25 (0.05 to 0.44)	0.01
Death	0.30 (0.10 to 0.50)	0.004	0.26 (0.01 to 0.51)	0.04

Adjusted analysis account for CXR score including cavitation, smear grade, ever smoking, diabetes and pre-treatment illness duration. Age, sex and BMI are adjusted by study design. Diabetes-restricted analysis are further adjusted for HbA1c levels. CI – confidence interval.

**Figure 4.** Difference in IL-6 concentrations between tuberculosis patients with and without unfavorable treatment outcomes, and associated area under the curve (AUC) among all participants (Panel A) and restricted to HIV coinfected patients only (Panel B) in the South African external validation cohort.





**Table 3.** Difference in IL-6 concentrations between with tuberculosis patients with and without unfavorable treatment outcomes after adjusting for markers of disease severity in the South African external validation cohort.

Treatment outcomes	Log <sub>10</sub> difference in baseline cytokine levels comparing participants with unfavorable treatment outcomes to cures			
	<i>Unadjusted difference (95%CI)</i>	<i>p-value</i>	<i>Adjusted difference (95%CI)</i>	<i>p-value</i>
<b>All participants</b>				
Composite	0.33 (0.06 to 0.60)	0.01	0.38 (0.13 to 0.63)	0.003
Failure	0.41 (0.05 to 0.76)	0.02	0.49 (0.15 to 0.82)	0.004
Recurrence	0.10 (-0.43 to 0.64)	0.70	-0.08 (-0.60 to 0.42)	0.73
Death	0.38 (-0.01 to 0.86)	0.10	0.50 (0.06 to 0.94)	0.02
<b>HIV coinfectd only</b>				
Composite	0.22 (-0.10 to 0.55)	0.15	0.29 (-0.01 to 0.60)	0.06
Failure	0.39 (-0.04 to 0.83)	0.06	0.46 (0.01 to 0.90)	0.04
Recurrence	-0.64 (-1.39 to 0.10)	0.10	-0.56 (-1.23 to 0.12)	0.10
Death	0.36 (-0.12 to 0.85)	0.13	0.48 (0.05 to 0.92)	0.03

Adjusted analyses account for age, sex, BMI, cavitation on chest X-ray, smear grade and HIV. HIV-restricted analyses are further adjusted for CD4 cell count and ART receipt. CI – confidence interval.

**Table 4.** Association between high concentration of IL-6 at baseline and unfavorable tuberculosis treatment outcomes in the pooled validation analysis.

<b>Treatment outcomes</b>	<b>Odds ratios for unfavorable treatment outcomes</b>			
	<i>Unadjusted (95%CI)</i>	<i>p-value</i>	<i>Adjusted (95%CI)</i>	<i>p-value</i>
Composite	3.32 (2.02-5.46)	<0.001	3.45 (2.03-5.89)	<0.001
Failure	2.16 (1.08-4.33)	0.02	2.22 (1.04-4.75)	0.03
Recurrence	5.36 (2.48-11.57)	<0.001	5.92 (2.52-13.90)	<0.001
Death	4.62 (1.95-10.95)	<0.001	5.53 (2.08-14.68)	<0.001

Regression analyses are adjusted for age, sex, BMI, cavitation on CXR, smear grade, HIV and diabetes.

## Supplementary Appendix 1

### Study population

**Discovery cohort:** The discovery cohort was designed for an exploratory analysis and its sample size was determined by project feasibility. Using a random number generator, we selected a sub-cohort of adults ( $\geq 18$  years) with newly diagnosed drug-sensitive pulmonary tuberculosis at the time of treatment initiation from the CTRIUMPH study at the Byramjee-Jeejeebhoy Government Medical College (BJGMC) in Pune(1). Cases were diagnosed by the presence of acid-fast bacilli (AFB) on smear microscopy, *Mycobacterium tuberculosis* DNA on Xpert MTB/RIF assay, *M tuberculosis* growth on culture, or based on clinical judgement in the absence of microbiological evidence of tuberculosis. HIV status was ascertained by ELISA (Genetic Systems HIV-1/HIV-2 PLUS O EIA, BioRad, USA). Diabetes was classified as having a prior physician diagnosis or a glycated hemoglobin (HbA1c) concentration  $\geq 6.5\%$  at enrollment. Chest X-ray (CXR) images were read by two independent reviewers and scored using a previously validated system which explicitly accounts for cavitation(2). Study participants received standard multi-drug tuberculosis treatment according to the Indian National Tuberculosis Elimination Program (NTEP) guidelines and, were followed for 24 months after treatment initiation to ascertain treatment outcomes of failure, recurrence or death. Plasma samples collected at treatment initiation, 2 months and 6 months underwent cytokine testing, in duplicates, using multiplex ELISA on Luminex assay (BioRad, USA) at the National Institutes of Health (NIH) – National Institute for Research in Tuberculosis (NIRT) – International Center for Excellence in Research (ICER) laboratory in Chennai. Cytokines analyzed were selected a-priori for their role in the

host inflammatory response to *M tuberculosis* (interferon gamma [INF- $\gamma$ ], tumor necrosis factor alpha [TNF- $\alpha$ ], interleukin [IL]-1 $\beta$ , IL-4, IL-6, IL-10, IP-10, IL-12, IL-13 and IL-17)(3), lung pathology (matrix metalloproteinases [MMP]-1, MMP-3, MMP-7, tissue inhibitor of metalloprotease [TIMP]-1, TIMP-2, TIMP-3, TIMP-4)(4), and fibrous remodeling (transforming growth factor beta [TGF $\beta$ ]-1, TGF $\beta$ -2, TGF $\beta$ -3)(5).

**Internal validation cohort:** We nested a case-control study within the CTRIUMPH cohort to validate statistically significant results from the discovery analysis. Adults with culture confirmed drug-sensitive pulmonary tuberculosis were eligible for selection; those already included in the discovery analysis were excluded from internal validation. Cases comprised of pulmonary tuberculosis patients who failed treatment, defined as confirmation of *M tuberculosis* by culture during the last two months of treatment. Controls comprised of tuberculosis patients with two consecutive cultures negative for *M tuberculosis* during the last two months of treatment. Controls were matched to cases on age and sex in a 1:1 ratio. Socio-demographic, clinical and laboratory data were collected using standardized questionnaires and protocols as described in the discovery cohort. Study participants received standard multi-drug tuberculosis treatment according to NTEP guidelines. IL-6, IL-13 and IFN- $\gamma$  measured at treatment initiation were statistically significantly associated with treatment failure in the discovery analysis and were therefore selected for internal validation. Plasma samples collected at treatment initiation underwent cytokine testing, in duplicates, using multiplex ELISA by Luminex assay (Bio-Rad, USA) at the BJGMC laboratory in Pune.

**Indian external validation cohort:** We nested a case-control study among adults with drug-sensitive pulmonary tuberculosis enrolled in the EDOTS cohort in Chennai(6).

Pulmonary tuberculosis was diagnosed by the presence of AFB on smear microscopy or culture confirmation of *M tuberculosis*; those with HIV coinfection were excluded. Participants received standard multi-drug tuberculosis treatment according to NTEP guidelines and were followed for 18 months after treatment initiation to ascertain outcomes. Cases comprised of participants with an unfavorable treatment outcome of failure, recurrence or death. Controls included participants with recurrence free cure during 18 months of follow-up. Controls were matched to cases on age, sex and BMI in a 2:1 ratio, with replacement. For participants reporting a prior history of diabetes, classification was confirmed by medical treatment history and HbA1c  $\geq$  6.5%. For participants with no known history of diabetes, classification was made by 75-gram oral glucose tolerance test as having a two-hour post-challenge blood glucose level  $\geq$  200mg/dL. CXR images were read by two independent reviewers and scored using a previously validated system which explicitly accounts for cavitation(2). IL-6 measured at treatment initiation was statistically significantly associated with treatment failure during internal validation and was therefore selected for independent external validation. Plasma samples collected at treatment initiation underwent IL-6 testing by Luminex assay (R&D Systems, USA) at the NIH-NIRT-ICER laboratory in Chennai.

**South African external validation cohort:** We nested a case-control study among adults with Xpert MTB/RIF-confirmed rifampin-susceptible pulmonary tuberculosis who were enrolled as part of a prospective cohort study in Khayelitsha, South Africa(7, 8). Participants received standard multi-drug tuberculosis treatment according to the South African National Tuberculosis Control Program guidelines and were followed for 22 months after treatment initiation. Cases comprised of participants with an unfavorable

treatment outcome of failure, recurrence or death. Controls included participants with recurrence free cure during 18 months of follow-up. HIV testing (Abbott Architect HIV Ag/Ab Combo test) and CD4+ T lymphocyte quantification were performed at enrollment. Diabetes was classified by a prior physician diagnosis or a HbA1c  $\geq$  6.5%. Plasma samples collected at tuberculosis treatment initiation underwent IL-6 testing by Luminex assay (MilliporeSigma, USA) at the Wellcome Centre for Infectious Disease Research, South Africa.

## References

1. Gupte A, Padmapriyadarsini C, Mave V, Kadam D, Suryavanshi N, Shivakumar SV, et al. Cohort for Tuberculosis Research by the Indo-US Medical Partnership (CTRIUMPH): protocol for a multicentric prospective observational study. *BMJ open*. 2016;6(2):e010542.
2. Ralph AP, Ardian M, Wiguna A, Maguire GP, Becker NG, Drogumuller G, et al. A simple, valid, numerical score for grading chest x-ray severity in adult smear-positive pulmonary tuberculosis. *Thorax*. 2010;65(10):863-9.
3. O'Garra A, Redford PS, McNab FW, Bloom CI, Wilkinson RJ, Berry MP. The immune response in tuberculosis. *Annu Rev Immunol*. 2013;31:475-527.
4. Ong CW, Elkington PT, Friedland JS. Tuberculosis, pulmonary cavitation, and matrix metalloproteinases. *American journal of respiratory and critical care medicine*. 2014;190(1):9-18.
5. Tatler AL, Jenkins G. TGF-beta activation and lung fibrosis. *Proc Am Thorac Soc*. 2012;9(3):130-6.
6. Kornfeld H, West K, Kane K, Kumpatla S, Zacharias RR, Martinez-Balzano C, et al. High Prevalence and Heterogeneity of Diabetes in Patients With TB in South India: A Report from the Effects of Diabetes on Tuberculosis Severity (EDOTS) Study. *Chest*. 2016;149(6):1501-8.
7. Rockwood N, Costa DL, Amaral EP, Du Bruyn E, Kubler A, Gil-Santana L, et al. Mycobacterium tuberculosis Induction of Heme Oxygenase-1 Expression Is Dependent on Oxidative Stress and Reflects Treatment Outcomes. *Front Immunol*. 2017;8:542.

8. Rockwood N, Sirgel F, Streicher E, Warren R, Meintjes G, Wilkinson RJ. Low Frequency of Acquired Isoniazid and Rifampicin Resistance in Rifampicin-Susceptible Pulmonary Tuberculosis in a Setting of High HIV-1 Infection and Tuberculosis Coprevalence. *The Journal of infectious diseases*. 2017;216(6):632-40.



## Supplementary Appendix 2

**Supplementary table 1.** Study schema for discovery, internal and external validation analyses.

	<b>Discovery</b>	<b>Internal validation</b>	<b>External validation 1</b>	<b>External validation 2</b>	<b>Pooled validation</b>
<b>Setting</b>	Pune, India	Pune, India	Chennai, India	Khayelitsha, South Africa	India and South Africa
<b>Design</b>	Prospective cohort study	Case-control analysis	Case-control analysis	Case-control analysis	Case-control analysis
<b>Sample size</b>	30 adults with drug-sensitive pulmonary tuberculosis	40 adults with drug-sensitive pulmonary tuberculosis	194 adults with drug-sensitive pulmonary tuberculosis	129 adults with rifampicin sensitive pulmonary tuberculosis	363 adults with drug-sensitive pulmonary tuberculosis
<b>Treatment outcomes</b>	4 with failure	20 with failure	18 with failure, 35 with recurrence and 19 deaths	9 with failure, 4 with recurrence and 5 deaths	47 with failure, 39 with recurrence and 24 deaths
<b>HIV coinfection, n (%)</b>	2 (7%)	3 (8%)	0	76 (59%)	79 (22%)
<b>Diabetes comorbidity, n (%)</b>	7 (23%)	3 (8%)	115 (59%)	10 (8%)	128 (35%)
<b>Laboratory</b>	NIH-NIRT-ICER, Chennai	BJGMC, Pune	NIH-NIRT-ICER, Chennai	Wellcome Center, Cape Town	BJGMC, NIH-NIRT-ICER and Wellcome Center

**Supplementary table 2.** Baseline characteristics and treatment outcomes of participants in the full cohort and those selected in the random sub-cohort for discovery analysis.

<b>Characteristics</b>	<b>Full cohort (n=445)</b>	<b>Sub-cohort (n=30)</b>	<b>p- value</b>
Age in years, median (IQR)	38 (27-49)	35 (28-51)	0.95
Male sex, n (%)	286 (64)	21 (70)	0.69
BMI in kg/m <sup>2</sup> , median (IQR)	17.7 (15.8-20.4)	18.0 (15.8-19.9)	0.97
Ever-smokers, n (%)	144 (32)	8 (27)	0.68
HIV coinfection, n (%)	33 (8)	2 (7)	0.99
Diabetes, n (%)	99 (22)	7 (23)	0.82
Pre-treatment illness duration in days, median (IQR)	45 (30-90)	60 (20-90)	0.59
Percent of lung fields affected on CXR, median (IQR)	60 (25-80)	50 (24-68)	0.30
Cavitation on CXR, n (%)	167 (45)	10 (40)	0.68
Culture confirmed TB, n (%)	368 (84)	27 (90)	0.45
Smear positive TB, n (%)	305 (69)	20 (67)	0.84
Treatment outcomes, n (%)			
Failure	45 (10)	4 (13)	0.99
Recurrence	22 (5)	0	-
Death	20 (4)	0	-

TB – tuberculosis, IQR – interquartile range, BMI – body mass index, CXR – chest radiograph, n – sample size.

**Supplementary table 3.** Baseline characteristics of culture confirmed tuberculosis patients who failed treatment (cases) and those who were cured (controls) in the internal validation cohort.

<b>Characteristics</b>	<b>Cases (n=20)</b>	<b>Controls (n=20)</b>	<b>p- value</b>
Age in years, median (IQR)	27 (20-32)	26 (20-32)	0.92
Male sex, n (%)	15 (75)	15 (75)	0.99
BMI in kg/m <sup>2</sup> , median (IQR)	16.0 (14.2- 17.7)	17.4 (15.8- 19.1)	0.02
Ever-smokers, n (%)	6 (30)	5 (25)	0.99
Alcohol dependence, n (%)	9 (45)	7 (35)	0.74
HIV coinfection, n (%)	2 (10)	1 (5)	0.99
Diabetes, n (%)	1 (5)	2 (10)	0.99
Pre-treatment illness duration in days, median (IQR)	35 (28-68)	30 (15-45)	0.03
Percent of lung fields affected on CXR, median (IQR)	58 (38-78)	60 (23-73)	0.43
Cavitation on CXR, n (%)	8 (40)	8 (40)	0.99
Smear positive TB, n (%)	14 (70)	11 (55)	0.51

TB – tuberculosis, IQR – interquartile range, BMI – body mass index, CXR – chest radiograph, n – sample size.

**Supplementary table 4.** Baseline characteristics of tuberculosis patients who experienced an unfavorable treatment outcome (cases) and those who were cured (controls) in the Indian external validation cohort.

<b>Characteristics</b>	<b>Cases (n=72)</b>	<b>Controls (n=122)</b>	<b>p-value</b>
Age in years, median (IQR)	45 (39-52)	45 (36-50)	0.40
Male sex, n (%)	64 (89)	104 (85)	0.52
BMI in kg/m <sup>2</sup> , median (IQR)	16.8 (15.2-18.9)	17.2 (15.4-19.2)	0.58
Ever-smokers, n (%)	41 (57)	50 (41)	0.03
Diabetes, n (%)	42 (58)	73 (59)	0.88
Pre-treatment illness duration in days, median (IQR)	4 (4-8)	5 (4-8)	0.56
Percent of lung fields affected on CXR, median (IQR)	26 (15-40)	24 (14-40)	0.36
Cavitation on CXR, n (%)	18 (31)	30 (32)	0.99
Smear positive TB, n (%)	61 (91)	94 (90)	0.80
Culture confirmed TB, n (%)	64 (96)	104 (98)	0.40

TB – tuberculosis, IQR – interquartile range, BMI – body mass index, CXR – chest radiograph, n – sample size.

**Supplementary table 5.** Baseline characteristics of TB patients who experienced an unfavorable treatment outcome and those who were cured in the South African external validation cohort.

<b>Characteristics</b>	<b>Unfavorable outcome (n=18)</b>	<b>Cure (n=111)</b>	<b>p-value</b>
Age in years, median (IQR)	36 (30-48)	35 (30-43)	0.78
Male sex, n (%)	11 (61)	63 (57)	0.80
BMI in kg/m <sup>2</sup> , median (IQR)	19.9 (19.0-23.0)	21.0 (19.1-23.3)	0.15
Ever-smokers, n (%)	8 (44)	52 (47)	0.99
Cavitation on CXR, n (%)	7 (39)	60 (54)	0.31
HIV coinfection, n (%)	13 (72)	63 (57)	0.30
ART receipt, n (%)	7 (39)	22 (20)	0.12
CD4 count, median (IQR)	178 (61-594)	194 (69-364)	0.98
Smear positive TB, n (%)	9 (50)	82 (74)	0.05

TB – tuberculosis, IQR – interquartile range, BMI – body mass index, CXR – chest radiograph, n – sample size, ART – antiretroviral therapy.

**Supplementary table 6.** Discriminatory ability of baseline IL-6 added to a risk-prediction model for unfavorable tuberculosis treatment outcomes.

<b>Baseline variables in the prediction model</b>	<b>AUC (95%CI) excluding baseline IL-6</b>	<b>AUC (95%CI) including baseline IL-6</b>	<b>p-value for gain in AUC by baseline IL-6</b>
Cavitation on CXR	0.52 (0.46-0.57)	0.68 (0.62-0.74)	<0.001
Smear grade $\geq 2$	0.50 (0.45-0.55)	0.65 (0.59-0.71)	<0.001
BMI $\leq 18.5$ kg/m <sup>2</sup>	0.60 (0.51-0.70)	0.72 (0.62-0.82)	0.002
Cavitation and smear grade $\geq 2$ and BMI $\leq 18.5$ kg/m <sup>2</sup>	0.66 (0.56-0.77)	0.76 (0.67-0.85)	0.02

AUC – area under the curve calculated by the C-statistic, CI – confidence interval, BMI – body mass index.