**Impact of intermediate hyperglycaemia as well as diabetes on immune dysfunction in tuberculosis**

Clare Eckold1, Vinod Kumar2,3, January Weiner 3rd4,Bachti Alisjahbana5, 5a, Anca-Lelia Riza3,6,6a, Katharina Ronacher7,8, Jorge Coronel9, Sarah Kerry-Barnard10, Stephanus T. Malherbe8, Leanie Kleynhans8, Kim Stanley8, Rovina Ruslami5, Mihai Ioana6,6a, Cesar Ugarte-Gil1,11, Gerhard Walzl8, Reinout van Crevel3, Cisca Wijmenga2, Julia A Critchley10, Hazel M. Dockrell1,\*, Jacqueline M. Cliff1,\*, on behalf of the TANDEM Consortium#

1TB Centre and Dept of Immunology and Infection, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, United Kingdom

2University of Groningen, University Medical Center Groningen, Department of Genetics, Groningen, the Netherlands

3Department of Internal Medicine and Radboud Center for Infectious Diseases, Radboud University Medical Center, Nijmegen, the Netherlands

4 Max Plank Institute for Infection Biology, Berlin, Germany

5TB-HIV Research Center, Faculty of Medicine, Universitas Padjadjaran, Bandung, Indonesia

5aHasan Sadikin General Hospital, Bandung, Indonesia

6Human Genomics Laboratory, University of Medicine and Pharmacy of Craiova, Romania

6aRegional Centre for Human Genetics – Dolj, Emergency Clinical County Hospital Craiova, Romania

7 Mater Research Institute – The University of Queensland, Translational Research Institute, Brisbane, Australia

8 SA MRC Centre for TB Research, DST/NRF Centre of Excellence for Biomedical Tuberculosis Research, Division of Molecular Biology and Human Genetics, Faculty of Medicine and Health Sciences, Department of Biomedical Sciences, Stellenbosch University, Cape Town, South Africa

9 Laboratorio de Investigación de Enfermedades Infecciosas, Universidad Peruana Cayetano Heredia, Lima, Peru

10 Population Health Research Institute, St George’s, University of London

11 School of Medicine, Universidad Peruana Cayetano Heredia, Lima, Peru

Short title: Blood transcriptomes in tuberculosis patients with diabetes

\* These authors contributed equally to the work.

#The full consortium author list is shown in the supplementary file

Corresponding author:

Dr Jackie Cliff,

Address: Dept of Infection Biology, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, United Kingdom,

Tel: +44 (0)207 927 2590

Email: jackie.cliff@lshtm.ac.uk

**Summary**

Our study shows that people with diabetes and tuberculosis have altered peripheral immune profiles compared to people with only tuberculosis, and further, these differences exist in patients with intermediate hyperglycaemia.

**Key words**

Tuberculosis, Diabetes, Inflammation

**Notes**

The research leading to these results, as part of the TANDEM Consortium, has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under Grant Agreement no. 305279.

Conflicts of Interest Statement: All authors declare they have no conflicts of interest.

**Abstract**

**Background**

People living with diabetes have an increased risk of developing active tuberculosis and are more likely to have poor tuberculosis-treatment outcomes, which may impact on control of tuberculosis as the prevalence of diabetes is increasing worldwide. Blood transcriptomes are altered in active tuberculosis patients relative to healthy individuals. The effects of diabetes and intermediate hyperglycaemia on this transcriptomic signature were investigated to enhance understanding of immunological susceptibility in diabetes-tuberculosis comorbidity.

**Methods**

Whole blood samples were collected from active tuberculosis patients with diabetes (HbA1c ≥6.5%) or intermediate hyperglycaemia (HbA1c 5.7-6.5%), tuberculosis-only patients and healthy controls in four countries: South Africa, Romania, Indonesia and Peru. Differential blood gene expression was determined by RNA-seq (n=249).

**Results**

Diabetes increased the magnitude of gene expression change in the host transcriptome in tuberculosis, notably showing an increase in genes associated with innate inflammatory and decrease in adaptive immune responses. Strikingly, patients with intermediate hyperglycaemia and tuberculosis exhibited blood transcriptomes much more similar to diabetes-tuberculosis patients than to patients with only tuberculosis. Both diabetes-tuberculosis and intermediate hyperglycaemia-tuberculosis patients had a decreased type I interferon response relative to tuberculosis-only patients.

**Conclusions**

Co-morbidity in individuals with both tuberculosis and diabetes is associated with altered transcriptomes, with an expected enhanced inflammation in the presence of both conditions, but also reduced type 1 interferon responses in co-morbid patients, suggesting an unexpected uncoupling of the TB transcriptome phenotype. These immunological dysfunctions are also present in individuals with intermediate hyperglycaemia, showing that altered immunity to tuberculosis may also be present in this group. The TB disease outcomes in individuals with intermediate hyperglycaemia diagnosed with TB should be investigated further.

**Introduction**

Despite much research, precise mechanisms underlying how the host immune response controls *Mycobacterium tuberculosis* are still not fully understood. Studies of comorbidity and coinfection associated with susceptibility to tuberculosis may provide useful insights. The epidemiological link between diabetes and tuberculosis (TB) has been known for centuries, yet the biological mechanisms underlying the association are still unclear [1]. People living with diabetes are over three times more likely to develop active TB once infected [2], and are also more likely to have poor TB-treatment outcomes [3]. Ninety percent of diabetes cases worldwide are Type 2 diabetes (T2DM) [4], characterised by insulin resistance and subsequent insulin insufficiency. Diabetes prevalence has risen from 108 million in 1980 to 415 million in 2014, with an estimated 642 million cases projected by 2040 [4]. The countries with the greatest number of people with diabetes include several countries with concurrent high TB burden, such as China, India, Brazil, Indonesia and Bangladesh. The estimated population attributable fraction of diabetes for adult TB is estimated at 9-17% across WHO regions [5], surpassing that for HIV in most regions, emphasising the critical need to target this population for TB control.

T2DM is characterised by a decrease in the ability of insulin to alter metabolism in target cells, leading to over-production of insulin, pancreatic β-cell exhaustion and eventual depletion of insulin, causing impaired glucose tolerance. Excessive glucose has toxic effects systemically, on major organs as well as on cells in the immune system. Additionally, T2DM is linked with obesity, with an increased number of adipocytes and with increased secretion of free fatty acids and of pro-inflammatory cytokines, including TNFα, IL-1β and IL-6 [6, 7], from adipocytes. In turn, these abnormalities in glucose control, dyslipidaemia and the chronic inflammatory response lead to an immune dysfunction disorder, affecting the body’s defence to *M. tuberculosis*. People with T2DM have impaired immune cell function, with reduced phagocytic ability in response to *M. tuberculosis,* which is correlated with poor glycaemic control [8]. Alongside chronic inflammation, including elevated circulating pro-inflammatory cytokines, people with T2DM-TB comorbidity have enhanced type 1 [9, 10] and regulatory T cell responses [11] to *M. tuberculosis* antigens, reviewed in [12].

Spectrums exist in both TB and T2DM, from latent *M. tuberculosis* infection through subclinical incipient TB to active TB disease [13], and from non-diabetes to clinically manifest diabetes via intermediate hyperglycaemia. Although not completely sensitive, glycated haemoglobin (HbA1c) measurement can determine an individual’s point along this spectrum, with the current threshold for diabetes diagnosis at HbA1c ≥ 6.5%, and for intermediate hyperglycaemia at 5.7% < HbA1c < 6.5% [14]. Intermediate hyperglycaemia has become an area of interest because it has also been shown to be associated with TB [15, 16].

Blood transcriptomic changes in TB patients have been well described: active TB patients are distinct from healthy controls, exhibiting a neutrophil derived interferon [17] and inflammatory response [18] signature. This resolves during successful TB drug treatment [19, 20], and the blood transcriptome in TB is distinct from that in other diseases [21, 22]. Blood-based transcriptomic studies of people living with diabetes are scarce. A recent study [23] found few differences in transcriptomes between TB-only patients and those with T2DM and TB comorbidity, although TB patients with intermediate hyperglycaemia were not defined in this study.

 Here, we used an unbiased RNA-seq approach to identify immunological pathways which were altered in T2DM-TB comorbidity compared to TB-only, and also investigated the impact of intermediate hyperglycaemia on TB, in order to understand why diabetes increases susceptibility to tuberculosis, and whether individuals with intermediate hyperglycaemia might also be at increased risk.

**Methods**

*Patient recruitment*

Newly diagnosed, bacteriologically confirmed, adult pulmonary TB patients with and without diabetes were recruited between December 2013 and February 2016, at four different study sites; Cape Town, South Africa (SUN), Bandung, Indonesia (UNPAD), Lima, Peru (UPCH) and Craiova, Romania (UMFCV). In South Africa and Romania, people with DM but without TB, and healthy individuals were also recruited. Patients were excluded if they were already taking TB treatment, had MDR-TB, were HIV-positive, pregnant, taking corticosteroids, or had other serious comorbidity. Diabetes (DM) was classified as laboratory test of glycated haemoglobin (HbA1c) ≥6.5%, alongside a confirmatory HbA1c ≥6.5% or fasting blood glucose ≥ 7mmol/L, as described in this cohort [24]. Intermediate hyperglycaemia (IH) patients were characterised with an HbA1c reading of >5.7% and <6.5%. Patients without diabetes, including healthy controls, had HbA1c values of <5.7%. All patients gave written informed consent, and the study was approved by the London School of Hygiene & Tropical Medicine (LSHTM) Observational/Interventions Research Ethics Committee (6449, 11/07/2013) as well as SUN Health Research Ethics Committee (N13/05/064, 29/07/2013), UNPAD Health Research Ethics Committee, Faculty of Medicine, Universitas Padjadjaran (No:377/UN6.C2.1.2/ KEPK/ PN/2012), UMFCV Committee of Ethics and Academic and Scientific Deontology (94, 06/09/2013), and UPCH Institucional Committee Of Ethics In Investigation (61069, 02/09/2013).

To limit confounding, prior to conducting the RNA-seq analysis, patient characteristic variables (age, body mass index, HbA1c) for each clinical group and for each study setting were initially tested using ANOVA. Subsequently, variables were tested for normality using the Shapiro-Wilk test:if deemed normally distributed, a pairwise t-test was performed, if not, a non-parametric Wilcoxon Rank Sum Test was performed. The participant ethnicities were recorded.

*RNA-seq experiments*

Venous blood was collected at TB diagnosis, before initiation of TB treatment, into PAXgene Blood RNA Tubes (PreAnalytiX) and used for RNA-seq analysis, using the polyA tail library preparation method and single-read sequencing (n=249). Sequencing FASTQ files were aligned to the human genome, and transcript quantification, differential gene expression, biological characterisation and machine learning were performed as described in the Supplementary Methods. RNA-seq quality control data are shown in Supplementary Figure 1 and Supplementary Table 1.

**Results**

*Study Population*

A total of 151 TB patients were recruited from four locations in South Africa, Romania, Indonesia and Peru, and classified as having TB-only, IH-TB or DM-TB, depending on lab HbA1c results at the time of TB diagnosis (Table 1). The DM-TB group included some patients with previously diagnosed DM, some of whom were taking insulin (Supplementary Table 2). Additionally, in South Africa and Romania, individuals without disease (healthy controls) and DM-only were recruited. The ages of the participants were not significantly different across the disease categories in any site or when the sites were combined (Table 1: Median age TB-only: 46 years, IH-TB: 46 years, DM-TB: 49 years). As anticipated, the BMI was significantly greater in the DM patients than in all the other patient groups at all sites (Table 1). Overall, participants had ten different stated ethnic backgrounds, but the participants’ ethnicities were not different across disease groups within each site.

*Blood TB transcriptome signature*

First, the effect of DM and IH on gene expression changes in TB patients’ blood was investigated in South African samples. As expected, there was substantial upregulation of gene expression (345 genes) in TB patients compared to healthy controls (Figure 1A), including upregulated expression of genes such as complement component C1QA,B,C and C2, BATF2, SOCS3, Septin 4, ANKRD22 and GBP5 which have been observed previously [20, 37, 38] (Figure 1A). Diabetes alone had a very different impact on the blood transcriptome, with larger numbers of genes differentially expressed, but with low magnitude of change and lower statistical significance (Figure 1B). In contrast, DM had a substantial impact on the blood transcriptome in TB patients, with 1,695 genes significantly up-regulated in DM-TB patients compared to healthy controls, including many ribosomal and transmembrane proteins, as well as pro-inflammatory cytokines such as IL-1β, IL-15 and IL-18, and the regulatory cytokine IL-10. There were also 1,623 genes significantly down-regulated in the DM-TB group compared to healthy controls, including several zinc finger transcription factors and cytokines such as IL-8, IL-16 and IL-24. Notably patients with IH-TB exhibited an even greater perturbation of their blood transcriptome relative to healthy controls, with 2,576 genes significantly up-regulated and 2,140 genes significantly down-regulated (Figure 1D). These results show that diabetes impacts the peripheral host response to TB, and that this effect is already present in individuals with intermediate hyperglycaemia. Supplementary Table 3 details all the differentially expressed genes in the South African cohort. Supplementary Figure 2 shows gender had no impact on gene expression differences.

These data were validated in a separate patient cohort, recruited in Romania. Similarly large scale up and down regulation of gene expression in both DM-TB and IH-TB was observed relative to healthy controls, to a much greater extent than that seen in TB-only (Supplementary Figure 3). Direct comparison of differentially expressed gene lists in South Africa and Romania showed significant overlap in each case (GeneOverlap P<0.05: Supplementary Figure 4).

LogFC-FC plots demonstrate concordant and discordant differential expression of genes when samples in different disease categories are compared to controls: discordant genes show significant differences in both comparisons, yet in opposite directions, while concordant genes show significant differences in both comparisons in the same direction. Here, the comparisons of all four disease groups with healthy controls in the South African cohort were compared (Figure 2). The main differences, showing the most discordance, were seen between the DM-only group vs HC comparison against all three comparisons of patients with TB (TB-only vs HC, DM-TB vs HC and IH-TB vs HC), showing that the TB transcriptome is dominant. The greatest concordance was between DM-TB vs HC and IH-TB vs HC, showing that patients in these categories exhibited similar changes in blood gene expression. The comparison between TB-only vs HC and IH-TB vs HC showed quantitative differences, with concordance between the genes but to a larger degree in the IH-TB group, indicating extra perturbation when hyperglycaemia is present (Figure 2).

To determine whether differential gene expression could be used to discriminate disease phenotype, Principal Component Analysis was performed on all study participants from South Africa. The resulting PCA plot, based on the differentially expressed genes from the initial analysis, showed two main clusters: healthy controls and people with DM-only clustered together, whilst the second main cluster comprised people with IH-TB and DM-TB comorbidity (Figure 3).

*Modular analysis in DM and TB comorbidity*

Modular analysis, which tests for enrichment sets of co-regulated genes sharing related biological functions [30, 31], was performed. As observed previously [17, 20], in the South African cohort, TB-only patients exhibited up-regulation of specific modules, including the type I interferon response, complement and activated dendritic cells (Figure 4; Supplementary Figure 5). In contrast, in diabetes patients, mainly respiration and protein synthesis modules were significantly up-regulated compared to healthy controls: these changes were not present in the comorbidity groups. Notably, module enrichments in DM-TB and IH-TB patients were similar to each other, which included up-regulation of modules differentially expressed in the TB-only group with a greater magnitude, especially for modules related to the inflammatory response and to myeloid cell function. There was stronger down-regulation of modules in IH-TB and DM-TB groups compared to TB-only: these included NK cell modules and adaptive immune response modules such as T-cell activation, differentiation and B-cells (Figure 4).

*Interferon and inflammatory responses*

To determine the reproducibility of results across geographical regions and different patient ethnicities, RNA-seq data from DM-TB, IH-TB and TB-only patients from all four study sites were combined. There was no systematic bias due to sample origin site (Supplementary Figure 6). In the combined dataset, DM-TB and IH-TB patients showed similar altered gene expression patterns in blood compared to patients with TB only (Figure 5). DM-TB patients had 292 upregulated and 130 downregulated genes, while IH-TB patients had 432 upregulated and 126 genes down-regulated (Supplementary Table 4).

In Modular analysis, both DM-TB and IH-TB showed a general trend of upregulation, especially of modules involved in inflammation. Interestingly however, patients with DM-TB and IH-TB also exhibited downregulation of several interferon modules compared to TB-only (Supplementary Figure 7). The most significantly differentially expressed are shown in Figure 6. The genes within these modules were differentially expressed in DM-TB and IH-TB, but of a greater magnitude in DM-TB, indicating an association with the extent of hyperglycaemia.

*Transcriptomic signatures*

We applied known; “Kaforou” [39] and “Sweeney” [40] TB biomarker signatures that could distinguish TB from latent TB infection, to test validity in DM patients, using the random forest algorithm. As expected, the Kaforou signature validated well in our TB-only dataset (AUC 0·96, Figure 7, blue line). However, in the DM-TB comorbidity cohort the performance was significantly reduced for both Kaforou (AUC=0.87, Figure 7, orange line, test for difference between ROC curves p=0.018) and Sweeney (AUC=0.84, Figure 7, green line) signatures. This implies that a TB biomarker signature for use in the general population needs to be derived from patient cohorts including those with DM. ROC curves showing sensitivities and specificities for these comparisons are shown in Supplementary Table 5.

**Discussion**

The principal finding of this study was that DM comorbidity influences blood transcriptomes in patients with TB, weakening the performance of published TB biosignatures. Of potentially greater public health importance, the modulation seen with diagnosed DM was also evident in those with intermediate hyperglycaemia, below the current HbA1c cut-off for DM diagnosis. There was greater upregulation of genes involved in the inflammatory response in diabetes and tuberculosis co-morbidity and also a reduced upregulation of the interferon response, principally of genes in the type 1 interferon pathway.

IH-TB patients were more similar to DM-TB patients than to people with TB-only, even though their hyperglycaemia was below the diabetes diagnostic threshold. Separation of the TB patients with HbA1c<6.5 facilitated the discovery of differences between the groups, which was not clearly evident in a previous transcriptomics study [23]. Our results demonstrate that patients with intermediate levels of hyperglycaemia exhibit immune dysfunction, which could lead to an increased susceptibility to active TB disease. It was previously proposed that adverse effects and disease susceptibility would occur at higher HbA1c values, later in diabetes progression [41]. Infectious diseases, including TB, can themselves induce transient hyperglycaemia [42] and poor outcome: further studies are warranted to establish the interplay between glycaemic control, TB and immune dysfunction. Targeted control of diabetes in TB could lead to better TB outcomes, and analysis of the direct impact of glycaemic control in immune dysfunction is merited.

Developing TB control strategies targeting people with diabetes is critical, due to the rapidly increasing global type 2 diabetes burden, especially in countries with high TB incidence [43]. People living with diabetes have an increased risk of developing active TB following infection [44], and therefore this patient group may contribute disproportionately to onward TB transmission. DM-TB comorbidity is more likely to result in poor TB treatment outcomes [3], including failure and relapse, adding an extra burden on healthcare systems. Globally, population level diabetes interventions might provide effective TB control strategies [45].

In a parallel TANDEM study analysis involving patients from the same four countries, TB patients with IH as well as with DM were more likely to be sputum smear positive compared to normoglycaemic TB controls [24]. The impact of intermediate hyperglycaemia on TB disease susceptibility has been less well investigated, but in two studies conducted in India [15, 46] and one in Kenya [47], both intermediate hyperglycaemia and diabetes prevalences in pulmonary TB patients were substantially higher than in the general population, indicating that intermediate hyperglycaemia is also associated with the development of active TB. Our data strongly indicate that immune abnormalities occurring in DM-TB are already present in IH-TB, and it is possible that existing altered immune function in hyperglycaemia underlies the epidemiological link.

This study shows that TB patients with either intermediate hyperglycaemia or diabetes have an altered immune phenotype, characterised by an excessive inflammatory response and a reduced type 1 IFN response. Both TB [18] and DM are pro-inflammatory conditions, and a degree of synergy between the two diseases is likely responsible for driving lung pathology and clinical symptoms. In active TB-only, type 1 interferon response genes are upregulated compared to healthy infected and uninfected individuals [17, 20, 22, 48]. Although surmised to be a deleterious excessive response in TB-only patients, the reduced IFN response seen in IH-TB and DM-TB patients might indicate an insufficient response, permitting continued growth of mycobacteria. Several interferon-related modules were present in the analysis performed, containing overlapping gene sets including genes encoding proteins involved in both type 1 and type 2 interferon pathways, all of which were all less up-regulated in the IH-TB and DM-TB groups compared to TB alone. The balance between inflammatory and type 1 IFN responses is critical for control of *M. tuberculosis* [49]*,* and our finding reveals a potential mechanism of TB susceptibility in intermediate hyperglycaemia as well as in diabetes patients. Further work is required to confirm whether the type 2 IFN response is also affected although we noted greater down-regulation of T cell and NK cell modules in IH-TB and DM-TB compared to TB alone. A recent study performed in Pakistan found raised serum concentrations of IFN and IL-13 in both IH-TB and DM-TB, although TB patients with up to 1 month of anti-tuberculosis therapy were included [50]. Further work could also include an analysis of regulatory miRNAs and lncRNAs, to understand better the mechanisms underpinning altered transcriptomes and immune responses in DM-TB.

Immunological differences in DM-TB compared to TB-only include enhanced pro-inflammatory responses [12], and increased circulating cytokines have also been described in IH-TB [51]. Immune abnormalities are detectable in individuals latently infected with *M. tuberculosis* with intermediate hyperglycaemia [52], and it is possible that concurrent intermediate hyperglycaemia and latent TB infection drives the development of both diseases, in a similar manner to the synergism between TB and HIV disease. This may indicate a relationship between moderate HbA1c values and TB susceptibility that warrants careful clinical considerations, and investigation of clinical disease outcome.

To conclude, these data have uncovered an enhanced inflammatory profile together with decreased interferon responses, associated with DM-TB comorbidity, which may be responsible for the increased susceptibility to TB in diabetes patients. Furthermore, immune dysfunction exists even at intermediate levels of hyperglycaemia, potentially causing TB susceptibility with implications for TB control. HbA1c values need to be considered in TB studies, as abnormal glycaemia clearly affects the immune response.

**Acknowledgments**

We thank the clinical staff and the patients at the four study sites. We thank Bahram Sanjabi, Desiree Brandenburg-Weening and Pieter van der Vlies for assistance with the RNA-seq.

**Members of the TANDEM Consortium**

The TANDEM partners and collaborators include: H Dockrell, J Cliff, C Eckold, D Moore, U Griffiths and Y Laurence from London School of Hygiene & Tropical Medicine, London, England; R Aarnouste, M Netea; R van Crevel, C Ruesen and E Lachmandas from Radboud University Medical Center, Nijmegen, Netherlands; S Kaufmann, M Beigier and R Golinski from Max Planck Institute for Infection Biology, Berlin Germany; S Joosten, T Ottenhoff, F Vrieling and M Haks from LUMC, Leiden, Netherlands; G Walzl, K Ronacher, S Malherbe, L Kleynhans, B Smith, K Stanley, G van der Spuy, A Loxton, N Chegou, M Bosman, L Thiart, C Wagman, H Tshivhula, M Selamolela, N Prins, W du Plessis, I van Rensburg and L du Toit from Stellenboch University, Cape Town, South Africa; J Critchley, S Kerry-Barnard, F Pearson, D Grint from St George’s, University of London, London, England; S McAllister, P Hill and A Verrall from University of Otago, Dunedin, New Zealand; M Ioana, A Riza, R Cioboata, M Dudau, F Nitu, I Bazavan, M Olteanu, C Editoiu, A Florescu, M Mota, SG Popa, A Firanescu, A Popa, I Gheonea, S Bicuti, A Lepadat, I Vladu, D Clenciu, M Bicu, C Streba, A Demetrian, M Ciurea, A Cimpoeru, A Ciocoiu, S Dorobantu, R Plesea, EL Popescu, M Cucu, I Streata, F Burada, S Serban-Sosoi, N Panduru and E Nicoli from Universtiy of Craiova, Craiova, Romania; M Ciontea, I Capitanescu, M Olaru, T Tataru, M Papurica, I Valutanu, V Dubreu, and L Stamatoiu from Pneumoftisiology Hospital, Gorj, Romania; V Kumar and C Wijmenga from University of Groningen, Groningen, Netherlands; C Ugarte-Gil, J Coronel, S Lopez, R Limascca, K Villaizan, B Castro, J Flores and W Solano from Universidad Peruana Cayetano, Lima, Peru; B Alisjahbana, R Ruslami, N Soetedjo, P Santoso, L Chaidir, R Koesoemadinata, N Susilawati, J Annisa, R Livia, V Yunivita, A Soeroto, H Permana, S Imaculata, Y Gunawan, N Dewi and L Apriani, Universitas Padjadjaran, Bandung, Indonesia.

**References**

1. Ronacher K, van Crevel R, Critchley JA, et al. Defining a research agenda to address the converging epidemics of tuberculosis and diabetes: Part 2: underlying biologic mechanisms. Chest **2017**; 152(1): 174-80.

2. Al-Rifai RH, Pearson F, Critchley JA, Abu-Raddad LJ. Association between diabetes mellitus and active tuberculosis: A systematic review and meta-analysis. PLoS One **2017**; 12(11): e0187967.

3. Baker MA, Harries AD, Jeon CY, et al. The impact of diabetes on tuberculosis treatment outcomes: a systematic review. BMC Med **2011**; 9: 81.

4. International Diabetes Federation. IDF Diabetes Atlas, **2017**.

5. Lonnroth K, Roglic G, Harries AD. Improving tuberculosis prevention and care through addressing the global diabetes epidemic: from evidence to policy and practice. The lancet Diabetes & endocrinology **2014**; 2(9): 730-9.

6. Spranger J, Kroke A, Mohlig M, et al. Inflammatory cytokines and the risk to develop type 2 diabetes: results of the prospective population-based European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam Study. Diabetes **2003**; 52(3): 812-7.

7. Kado S, Nagase T, Nagata N. Circulating levels of interleukin-6, its soluble receptor and interleukin-6/interleukin-6 receptor complexes in patients with type 2 diabetes mellitus. Acta Diabetol **1999**; 36(1-2): 67-72.

8. Restrepo BI, Twahirwa M, Rahbar MH, Schlesinger LS. Phagocytosis via complement or Fc-gamma receptors is compromised in monocytes from type 2 diabetes patients with chronic hyperglycemia. PLoS One **2014**; 9(3): e92977.

9. Kumar NP, Sridhar R, Banurekha VV, Jawahar MS, Nutman TB, Babu S. Expansion of pathogen-specific T-helper 1 and T-helper 17 cells in pulmonary tuberculosis with coincident type 2 diabetes mellitus. J Infect Dis **2013**; 208(5): 739-48.

10. Stalenhoef JE, Alisjahbana B, Nelwan EJ, et al. The role of interferon-gamma in the increased tuberculosis risk in type 2 diabetes mellitus. Eur J Clin Microbiol Infect Dis **2008**; 27(2): 97-103.

11. Sun Q, Zhang Q, Xiao H, Cui H, Su B. Significance of the frequency of CD4+CD25+CD127- T-cells in patients with pulmonary tuberculosis and diabetes mellitus. Respirology **2012**; 17(5): 876-82.

12. Ronacher K, Joosten SA, van Crevel R, Dockrell HM, Walzl G, Ottenhoff TH. Acquired immunodeficiencies and tuberculosis: focus on HIV/AIDS and diabetes mellitus. Immunol Rev **2015**; 264(1): 121-37.

13. Suliman S, Thompson E, Sutherland J, et al. Four-gene pan-African blood signature predicts progression to tuberculosis. Am J Respir Crit Care Med **2018**; 197(9): 1198–208.

14. American Diabetes A. Diagnosis and classification of diabetes mellitus. Diabetes Care **2010**; 33 Suppl 1: S62-9.

15. Viswanathan V, Kumpatla S, Aravindalochanan V, et al. Prevalence of diabetes and pre-diabetes and associated risk factors among tuberculosis patients in India. PLoS One **2012**; 7(7): e41367.

16. Critchley JA, Restrepo BI, Ronacher K, et al. Defining a research agenda to address the converging epidemics of tuberculosis and diabetes: Part 1: epidemiology and clinical management. Chest **2017**; 152(1): 165-73.

17. Berry MP, Graham CM, McNab FW, et al. An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. Nature **2010**; 466(7309): 973-7.

18. Joosten SA, Fletcher HA, Ottenhoff TH. A helicopter perspective on TB biomarkers: pathway and process based analysis of gene expression data provides new insight into TB pathogenesis. PLoS One **2013**; 8(9): e73230.

19. Bloom CI, Graham CM, Berry MP, et al. Detectable changes in the blood transcriptome are present after two weeks of antituberculosis therapy. PLoS One **2012**; 7(10): e46191.

20. Cliff JM, Lee JS, Constantinou N, et al. Distinct phases of blood gene expression pattern through tuberculosis treatment reflect modulation of the humoral immune response. J Infect Dis **2013**; 207(1): 18-29.

21. Bloom CI, Graham CM, Berry MP, et al. Transcriptional blood signatures distinguish pulmonary tuberculosis, pulmonary sarcoidosis, pneumonias and lung cancers. PLoS One **2013**; 8(8): e70630.

22. Maertzdorf J, Weiner J, 3rd, Mollenkopf HJ, et al. Common patterns and disease-related signatures in tuberculosis and sarcoidosis. Proc Natl Acad Sci U S A **2012**; 109(20): 7853-8.

23. Prada-Medina CA, Fukutani KF, Pavan Kumar N, et al. Systems Immunology of Diabetes-Tuberculosis Comorbidity Reveals Signatures of Disease Complications. Sci Rep **2017**; 7(1): 1999.

24. Ugarte-Gil C, Alisjahbana B, Ronacher K, et al. Diabetes mellitus among pulmonary tuberculosis patients from 4 tuberculosis-endemic countries: the TANDEM study. Clin Infect Dis **2020**; 70(5): 780-8.

25. Dobin A, Davis CA, Schlesinger F, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics **2013**; 29(1): 15-21.

26. Andrews S. FastQC: a quality control tool for high throughput sequence data. **2010**.Available at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>. Accessed April 2020.

27. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics **2015**; 31(2): 166-9.

28. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol **2014**; 15(12): 550.

29. Shen L. GeneOverlap: An R package to test and visualize gene overlaps. **2013**. Available at: <http://shenlab-sinai.github.io/shenlab-sinai/>. Accessed April 2020.

30. Chaussabel D, Quinn C, Shen J, et al. A modular analysis framework for blood genomics studies: application to systemic lupus erythematosus. Immunity **2008**; 29(1): 150-64.

31. Li S, Rouphael N, Duraisingham S, et al. Molecular signatures of antibody responses derived from a systems biology study of five human vaccines. Nat Immunol **2014**; 15(2): 195-204.

32. Weiner 3rd J, Domaszewska T. tmod: an R package for general and multivariate enrichment analysis. PeerJ Preprints **2016**; 4: e2420v1

33. Zyla J, Marczyk M, Domaszewska T, Kaufmann SHE, Polanska J, Weiner J. Gene set enrichment for reproducible science: comparison of CERNO and eight other algorithms. Bioinformatics **2019**; 35(24): 5146-54.

34. Benjamini Y., Y. H. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the royal statistical society Series B (Methodological) **1995**; 57: 289-300.

35. Domaszewska T, Scheuermann L, Hahnke K, et al. Concordant and discordant gene expression patterns in mouse strains identify best-fit animal model for human tuberculosis. Sci Rep **2017**; 7(1): 12094.

36. Breiman L. Random Forests. Machine Learning **2001**; 45: 5-32.

37. Roe JK, Thomas N, Gil E, et al. Blood transcriptomic diagnosis of pulmonary and extrapulmonary tuberculosis. JCI Insight **2016**; 1(16): e87238.

38. Zak DE, Penn-Nicholson A, Scriba TJ, et al. A blood RNA signature for tuberculosis disease risk: a prospective cohort study. Lancet **2016**; 387(10035): 2312-22.

39. Kaforou M, Wright VJ, Oni T, et al. Detection of tuberculosis in HIV-infected and -uninfected African adults using whole blood RNA expression signatures: a case-control study. PLoS Med **2013**; 10(10): e1001538.

40. Sweeney TE, Braviak L, Tato CM, Khatri P. Genome-wide expression for diagnosis of pulmonary tuberculosis: a multicohort analysis. The Lancet Respiratory Medicine **2016**; 4(3): 213-24.

41. Lee PH, Fu H, Lai TC, Chiang CY, Chan CC, Lin HH. Glycemic control and the risk of tuberculosis: A Cohort Study. PLoS Med **2016**; 13(8): e1002072.

42. Dungan KM, Braithwaite SS, Preiser JC. Stress hyperglycaemia. Lancet **2009**; 373(9677): 1798-807.

43. IUATLD/WHO: Collaborative framework for care and control of tuberculosis and diabetes. 2011. Available at: <http://whqlibdoc.who.int/publications/2011/9789241502252_eng.pdf>. Accessed April 2020.

44. Jeon CY, Murray MB. Diabetes mellitus increases the risk of active tuberculosis: a systematic review of 13 observational studies. PLoS Med **2008**; 5(7): e152.

45. Awad SF, Critchley JA, Abu-Raddad LJ. Epidemiological impact of targeted interventions for people with diabetes mellitus on tuberculosis transmission in India: Modelling based predictions. Epidemics **2019**; 30: 100381.

46. Mave V, Meshram S, Lokhande R, et al. Prevalence of dysglycemia and clinical presentation of pulmonary tuberculosis in Western India. Int J Tuberc Lung Dis **2017**; 21(12): 1280-7.

47. Owiti P, Keter A, Harries AD, et al. Diabetes and pre-diabetes in tuberculosis patients in western Kenya using point-of-care glycated haemoglobin. Public Health Action **2017**; 7(2): 147-54.

48. Ottenhoff TH, Dass RH, Yang N, et al. Genome-wide expression profiling identifies type 1 interferon response pathways in active tuberculosis. PLoS One **2012**; 7(9): e45839.

49. Mayer-Barber KD, Andrade BB, Oland SD, et al. Host-directed therapy of tuberculosis based on interleukin-1 and type I interferon crosstalk. Nature **2014**; 511(7507): 99-103.

50. Hasan Z, Irfan M, Masood Q, et al. Raised levels of IFN-gamma and IL-13 are associated with pre-diabetes amongst newly diagnosed patients with Tuberculosis. J Pak Med Assoc **2019**; 69(4): 468-73.

51. Kumar NP, Banurekha VV, Nair D, et al. Coincident pre-diabetes is associated with dysregulated cytokine responses in pulmonary tuberculosis. PLoS One **2014**; 9(11): e112108.

52. Kumar NP, Moideen K, Dolla C, Kumaran P, Babu S. Prediabetes is associated with the modulation of antigen-specific Th1/Tc1 and Th17/Tc17 responses in latent Mycobacterium tuberculosis infection. PLoS One **2017**; 12(5): e0178000.

**Table 1: Study participants’ characteristics**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | TB Only | IH-TB | DM-TB | T2DM only | Healthy Controls | P-value (ANOVA) |
| Number of study participants | S. Africa | 11 | 20 | 15 | 33 | 24 | - |
| Romania | 10 | 10 | 15 | 19 | 12 | - |
| Indonesia | 14 | 5 | 19 | - | - | - |
| Peru | 11 | 9 | 12 | - | - | - |
| All sites  | 46 | 44 | 61 | - | - | - |
| Age(Median, Range) | S. Africa | 48 (31-56) | 44·5 (25-57) | 46 (27-57) | 49 (29-64) | 42 (30-70) | 0·168 |
| Romania | 43 (30-64) | 48.5 (22-63) | 47 (22-64) | 55 (38-65) | 46 (38-61) | 0·329 |
| Indonesia | 47 (28-62) | 51 (37-54) | 52 (33-66) | - | - | 0·430 |
| Peru | 55 (31-69) | 52 (31-68) | 50·5 (42-58) | - | - | 0·928 |
| All sites  | 46 (28-69) | 46 (22-68) | 49 (22-66) | - | - | 0·2645 |
| Gender%Male (number male/number female) | S. Africa | 18 (2/9) | 60 (12/8) | 47 (7/8) | 45 (15/18) | 50 (12/12) | - |
| Romania | 60 (6/4) | 90 (9/1) | 87 (13/2) | 73 (14/5) | 83 (10/2) | - |
| Indonesia | 50 (7/7) | 80 (4/1) | 58 (11/8) | - | - | - |
| Peru | 45 (5/6) | 55 (5/4) | 50 (6/6) | - | - | - |
| All sites  | 43 (26/20) | 68 (30/14) | 60 (37/24) | - | - | - |
| Body Mass Index(Median, Range) | S. Africa | 20·1 (14·6 24·1) | 18·4 (13·7 -27·1) | 19·1 (13·9 -32·3) | 29 (20·52 -52·54) | 23·7 (17·37 45·20) | 1·43e-08 |
| Romania | 20·65 (17·7 -25·5) | 21 (15·5 22) | 21·5 (15·3 36·1) | 28·7 (13·1 -37·8) | - | 6·73e-04 |
| Indonesia | 19·8 (13·76 -33·27) | 18·67 (13·9-20·09) | 19·52 (16·10 31·73) | - | - | 0·242 |
| Peru | 23·98 (17·35 -28·96) | 21·94 (18·67 -25·56) | 22·7 (20·55 -33·33) | - | - | 0·207 |
| All sites  | 20·94 (13·76 33·27) | 19·44 (13·7 27·1) | 21·5 (13·9 36·1) | - | - | 1·67e-03 |
| HbA1c (%)(Median, Range) | S. Africa | 5·3 (4·8-5·7) | 6 (5·7-6·3) | 10·8 (6·5 -14·3) | 10·1 (4·7-14·9 | 5·3 (4·8-5·6) | 2·2e-16 |
| Romania | 5·5 (5·2-5·7) | 6 (5·7-6·4) | 10·4 (6·6-15) | 9 (7·1 14) | 5·4 (5·2-5·6) | 1·2e-13 |
| Indonesia | 5·55 (5-5·7) | 5·9 (5·8-6) | 11·9 (7·5 -15·9) | - | - | 3·20e-12 |
| Peru | 5·3 (5·1-5·7) | 5·8 (5·7-6·1) | 11·4 (7·1 -15·2) | - | - | 8·33e-09 |
| All sites  | 5·5 (4·8-5·7) | 5·95 (5·7-6·4) | 11 (6·5 15·9) | - | - | 2·2e-16 |

**Figure Legends**

**Figure 1: Differential expression analysis of all the disease phenotypes in South Africa compared to healthy controls before the initiation of TB treatment.** Gene expression profiles of A) TB-only (n=11); B) DM-only (n=33), C) DM-TB (n=15), D) IH-TB (n=20), each relative to healthy controls (n=24). Genes that were deemed statistically significantly differentially expressed had an adjusted P < 0.05 after multiple testing correction (Benjamini & Hochberg). Purple corresponds to the genes whose expression was significantly changed, grey shows genes without significant expression change. FDR: false Discovery Rate.

**Figure 2:** **Concordance and discordance of gene expression between the comparisons of each disease group and healthy controls in South Africa.** Log fold changes and p-values between groups was calculated with R-package DESeq2. A disco.score was calculated for each pair of corresponding genes. The axes show log2 fold change between the conditions indicated by the labels. For example, on the top left plot the X axis corresponds to the comparison between TB and HC, and the Y axis shows the log2 fold change between DM-TB and HC. Red dots show genes that are significantly different from the controls in the same direction (concordant genes), and blue dots show genes which are significantly different in both comparisons, but in opposite directions. Intensity of colour indicates the strength of concordance / discordance as measured by the disco.score.

**Figure 3: Principal Component Analysis of South African participants.** The list of all genes which were significantly differentially expressed in any patient group comparison with healthy controls was used in a Principal Component Analysis of all the samples obtained from participants recruited in South Africa.

**Figure 4: Transcriptional modules that were significantly differentially expressed in TB-only, DM-TB, IH-TB and DM-only compared to healthy controls in South Africa before initiation of TB treatment.** Transcripts were evaluated using a pre-existing modular framework. Significantly up-regulated (red) and down-regulated (blue) modules are shown: the length of each bar corresponds to the effect size (magnitude of change) of that module, and the colour saturation represents the adjusted p-value (< 0.0001). The amount of colour represents the proportion of genes within that module that were differentially expressed.

**Figure 5: Differential gene expression analysis of DM-TB and IH-TB patients relative to TB-only patients, in the combined dataset from all four field sites.** Samples collected in South Africa, Romania, Peru and Indonesia from A) DM-TB patients and B) IH- patients were compared with patients with TB-only in a combined analysis. Genes significantly differentially expressed after multiple testing correction are shown in pink (p-value <0.05). Genes in grey are not statistically significantly altered compared to patients with untreated TB only. FDR: false discovery rate.

**Figure 6: Summary of modular analysis in all four field sites.** The fold changes of the genes within the top significantly differentially expressed modules are shown (adjusted p-value < 0.05). On the inside: IH-TB compared to TB-only. Outside: DM-TB compared to TB-only. Up-regulated genes are shown in red, and down-regulated genes are in blue. The saturation of colour represents the magnitude of differential expression.

**Figure 7: Predictive model of known signature in TANDEM data**

Receiver operating characteristic curves based on a machine learning model generated from two different external data set of transcriptome profiles of TB patients and healthy controls (Kaforou [34]and Sweeney [40] training set). The random forest model was applied to the TANDEM cohort (test set), separately to individuals with (“DM”) and without DM (“no DM”).