Full Title (140 Characters): Soluble biomarkers associated with chronic lung disease in older children and adolescents with perinatal HIV infection

Short Title (40 Characters): Biomarkers with associated HCLD

Abstract Word Count: 240

Total Manuscript Word Count: 2382

Authors:

Dan Hameiri Bowen[1], MSc

Evgeniya Sovershaeva[2,3], PhD

Trond Flaegstad [2,3], PhD

Tore Jarl Gutteberg [2,3], PhD

Lucky Gift Ngwira[6], MPH

Victoria Simms[4] PhD

Andrea M Rehman[4] PhD

Grace McHugh[4,5], MD

Tsitsi Bandason[5], MSc

Rashida Abbas Ferrand [4,5] PhD

Sarah Rowland-Jones[1], DM

Louis-Marie Yindom[1], PhD*

Running title: Soluble Biomarkers in HCLD

1 University of Oxford, Nuffield Department of Medicine, Oxford, United Kingdom
2 UiT The Arctic University of Norway, Tromsø, Norway
3 University Hospital of North Norway, Tromsø, Norway
4 London School of Hygiene and Tropical Medicine, London, United Kingdom
5 Biomedical Research and Training Institute, Harare, Zimbabwe
6 Liverpool School of Tropical Medicine, Liverpool, UK & Malawi-Liverpool Wellcome Trust Clinical Research Program, Blantyre, Malawi

Keywords: HIV, Lung disease, Biomarkers, pathogenesis, Africa, Adolescents

Funding received for work: This work was supported by Research Council of Norway, through their GLOBVAC programme, with additional support from the regional hospital trust of North Norway, Bodo, Norway. RAF is funded by the Wellcome Trust through a Senior Fellowship in Clinical Science.

Correspondence:

Louis-Marie Yindom

NDM Research Building, Nuffield Department of Medicine, University of Oxford, Old Road Campus, Roosevelt Drive, Headington, Oxford, OX3 7F

Abstract (250 Words) – AIDS Style

***Objective**:* HIV-associated chronic lung disease (HCLD) is a common comorbidity in children and adolescents in sub-Saharan Africa (SSA). The pathogenesis of HCLD is unclear and may be driven by underlying dysregulated systemic immune activation and inflammation. We investigated the association between a wide spectrum of plasma soluble biomarkers and HCLD.

***Design**:* Case control analysis of baseline data from ART-treated participants recruited into the "Bronchopulmonary Function in Response to Azithromycin Treatment for Chronic Lung Disease in HIV-infected Children" (BREATHE) clinical trial.

***Methods**:* We recruited 6-19-year olds with perinatal HIV infection (PHIV) through HIV clinics in Malawi and Zimbabwe. Cases, with HCLD, had forced expiratory volume in one second (FEV_1) z-score <-1 with no reversibility. Controls, without HCLD, had FEV_1 z-score > 0.

Logistic regression identified biomarkers associated with the odds of HCLD. Data-reduction techniques were employed to find groupings of biomarkers associated with HCLD and FEV_1 z-score.

***Results**:* We recruited 336 cases and 74 controls. Biomarkers of general immune activation and inflammation (B2M, CRP, sCCL5, GCSF, IFN-γ, IP-10), T-Cell activation (sCD25, sCD27, sCD40-Ligand), monocyte activation (sCD14), coagulation (D-Dimer), cellular adhesion (E-selectin), and extracellular matrix degradation (MMP-1, MMP-7, MMP-10) were associated with increased odds of HCLD. Exploratory PCA identified a component comprising

predominantly T-cell activation markers (sCD40-Ligand, sCD25 and sCD27) that associated with increased odds of HCLD.

Conclusions: Biomarkers of inflammation, T-cell activation, monocyte activation, coagulation and extracellular matrix degradation were found to be associated with HCLD.

Introduction

The widespread use of combination antiretroviral therapy (ART) has led to a growing number of children and adolescents with perinatally-acquired HIV infection (PHIV) in sub-Saharan Africa (SSA) surviving into adolescence and beyond (1). In recent years, a range of chronic cardiovascular, respiratory, musculoskeletal and neurocognitive comorbidities have been described among children growing up with HIV, despite ART (2–5). In particular, while ART treatment has reduced the incidence of pulmonary infections, there remains a substantial burden of chronic respiratory symptoms along children and adolescents with HIV (6–8). Studies have reported a prevalence of about 30% in children with HIV aged over 10 years (2). HIV associated chronic lung disease (HCLD) is typically characterised by a chronic cough, exercise restriction, hypoxia and airflow obstruction without reversibility (9).

The underlying pathological processes contributing to HCLD are poorly understood. Immune activation and inflammation are key mechanisms in the pathogenesis of multiple chronic complications of adult HIV, and are associated with airflow obstruction in HIV-infected adults (10–12). The underlying mechanisms could be unique to the pediatric population or might be shared with adult HIV infection. Ongoing airway inflammation, either directly due

to HIV infection or infections that occur as a consequence of HIV-related immunosuppression may result in progressive tissue remodeling, fibrosis of the small airways and lung function decline (13).

We conducted a case control study to investigate the association of soluble biomarkers encompassing a wide spectrum of pathogenetic pathways with HCLD in children in Malawi and Zimbabwe.

Methods:

This study was nested within the BREATHE trial (**Br**onchopulmonary function in r**e**sponse to **a**zithromycin **t**reatment for chronic lung disease in **H**IV-infected children) (ClinicalTrials.gov, NCT02426112), which investigated the impact of azithromycin therapy on lung function in children with HCLD. Full details are described elsewhere (14–16). Briefly, inclusion criteria were age 6 to 19 years, perinatally-acquired HIV, taking ART for at least 6 months and HCLD (FEV_1 z-score <-1.0 with no reversibility on bronchodilators). Participants were recruited from two public sector HIV clinics in Harare, Zimbabwe and Blantyre, Malawi. Individuals with acute respiratory tract infections, tuberculosis or potentially fatal conditions at time of screening were excluded. Trial participants served as cases for the current study. Controls without HCLD (FEV_1 z-score >0) but otherwise meeting the same criteria and frequency-matched to cases by age and duration on ART were recruited for laboratory studies.

Measurement of Soluble Biomarkers

Biomarkers were measured from cryopreserved plasma stored at -80^{o}C. The full list of biomarkers measured are described in Table 1. The levels of all plasma soluble biomarkers were measured using the Luminex multiplex bead assay on a MagPix instrument according to the manufacturer's protocol (Luminex technology, Hertogenbosch, Netherlands). All plasma samples were used on their first thaw and were measured in duplicate on the same machine. Biomarker measurements falling outside of the standard curve were repeated at an adjusted dilution.

Statistical Methods

Data were analysed in R Studio (Version 1.1.383). For continuous demographic and anthropometric variables, the mean, median and interquartile range (IQR) were calculated by HCLD status. For categorical variables proportions were calculated. Differences between cases and controls were assessed by Kruskal Wallis test for continuous variables and Chi-square test for categorical variables. Weight-for-age and height-for-age z-scores were calculated using British 1990 Growth Reference Curves (17). Wasting and stunting were defined respectively as weight and height for age z-score less than -2.

Biomarkers with detectable levels falling below the standard curve were assigned half the minimum value measured. Spearman rank correlation coefficients between all soluble biomarkers were calculated in all participants and within cases, and correlation networks were visualised using the qgraph package in R. Statistically significant ($p < .05$) Spearman rank correlations were visualised. Within the case and control network the centrality (interconnectedness) of each biomarker was calculated and converted into a z-score to facilitate between biomarker comparisons. To increase comparability of regression results between biomarkers, all biomarkers were scaled to a mean of 0 and standard deviation of 1 within the population studied. Logistic regression was used to assess the association of biomarkers with HCLD. In cases, the association between biomarker and FEV_1 z-score as a continuous measure was assessed using linear regression. Variables associated with HCLD, FEV_1 z-score and biomarker level, alongside those defined *a priori* were included as covariates in adjusted models. Sex, trial site, supressed viral load (HIV viral load <200 copies/ml) and having ever been treated for TB were included as binary covariates. Age and height-for-age z-score were included as continuous variables. Adjusted and unadjusted odds ratios, 95% confidence intervals (CI), regression coefficients and their standard errors are presented. Where required, missing data in clinical covariates were imputed by mean imputation where appropriate.

Due to the expected correlation between biomarkers, techniques were employed to reduce the dimensionality of the data. Exploratory principal component analysis (PCA) was performed using the FactoMineR package in R (18). Prior to assessment, biomarker values were scaled. PCA dimensions with eigenvalues >1 were retained for downstream analysis.

Exploratory PCA was performed separately for all participants and then for cases. Participants value for each principal component were extracted and included in the logistic and linear regression analysis described. The sensitivity and specificity of biomarker levels for predicting HCLD was assessed using receiver operating characteristics (ROC) analysis. Area under the curve (AUC) of each biomarker and HCLD were calculated for all biomarkers and principal components showing association with HCLD. Threshold values maximising sensitivity and specificity for each biomarker were calculated.

Ethics

Consent from individuals within BREATHE study was sought from the guardian and age-appropriate assent from the participant (for those aged <18 years).

Results

A total of 410 participants (336 cases and 74 controls) were recruited. Cases were more likely than controls to be stunted (50.0% vs 29.7% p< .001), have ever been treated for TB (28.9% vs 12% p = .005) and to be on first line ART (26.2% vs 10.8%, p= .008) (Table 2).

Biomarker levels (pg/ml) are reported by group in Supplementary Table 1. A total of 59 (0.6%) biomarker measurements fell below the limit of detection and came from two biomarkers (MMP-12 (54) and Angiopoietin 1(5). These individuals were assigned half the

lowest measured value from the whole population. Owing to this, MMP-12 was dropped from all data-reduction techniques. One data point was missing for GDF-15 which was imputed by mean imputation. There was a high degree of correlation between all biomarkers (Figure 1). Network plots of all biomarkers split by HCLD identified CD40-Ligand as the most interconnected biomarker in cases and sCD25 in controls (Supplementary Figure 1A/B). The interconnectedness z-score of each individual biomarker is displayed in Supplementary Figure 1C. IFN-γ, IP-10, ANG-1, D-dimer and B2M all showed noticeable differences in interconnectedness between cases and controls.

Adjusted logistic regression identified 18 soluble biomarkers associated with HCLD status. 16 biomarkers were associated with increased the odds of HCLD (MMP-1, MMP-7, MMP-10, Angiopoietin-1, sCCL5, sCD14, sCD25, sCD27, sCD40-Ligand, CRP, IP-10, D-Dimer, E-Selectin, Fas, IFN-γ, VCAM-1) and two were associated with reduced odds (GCSF and VEGF) (Table 3).

Among cases, 10/26 biomarkers were associated with reduced FEV_1 z-score (MMP-8, MMP-10, ANG-1, CRP, IP-10, E-Selectin, Fas, GCSF, VCAM-1 and VEGF) (Table 3). The coefficients were largest for MMP-10 and CRP ($\beta = -0.132 \pm 0.04$ & $\beta = -0.128 \pm 0.38$). An increase of one standard deviation in Fas was associated with a small increase in FEV_1 z-score ($\beta = 0.083 \pm 0.039$, $p = .028$). Of these biomarkers, only MMP-8 was not associated with HCLD in the previous analysis.

Exploratory PCA of all participants identified seven components with eigenvalues >1, explaining 61% of the variance (Supplementary Table 4). The variance explained by each component and the contribution of biomarkers to each component are presented in Supplementary Figure 2A/B. Principal component one (predominantly comprising T-cell activation markers sCD40-Ligand and sCD25 alongside IFN-γ) was associated with significantly increased odds of HCLD (OR= 1.54 95% CI=1.33-1.80, p= < .001) (Supplementary Table 6). PCA within cases identified 8 dimensions with eigenvalues > 1 which explained 64% of the variation **(**Supplementary Table 4). The contribution of the top six biomarkers to each dimension is shown in Supplementary Figure 1B. Principal component one predominantly composed of immune activation markers sCD40-Ligand, B2M and IFN-γ, had evidence of association with reduced FEV_1 z-score (β = -0.04 (0.018), p= .014). FEV_1 z-score was also associated with dimensions three four and seven (Supplementary Table 6).

Receiver operator characteristics (ROC) identified CD40-Ligand, sCD25 and PCA principal component 1 as having area under the curve (AUC) greater than 0.7 (Supplementary Figure 3 & Supplementary Table 4). The best performing biomarker was log_{10} sCD40-Ligand which at a threshold of 3.526 had a specificity of 0.716, sensitivity of 0.812 and AUC 0.768.

Discussion

Soluble biomarkers have been associated with reduced lung function in multiple diseases (25,26), including HIV (19,20). Radiological findings from similar populations leads us to believe that reduced lung function in this population is caused by obliterative bronchiolitis

(OB) (2,3). OB is a condition characterized by inflammation and fibrosis of the terminal bronchioles resulting in progressive airflow obstruction and lung function decline (21–23). In this study, we describe an association between soluble biomarkers involved in several pathways and HCLD, suggesting potential mechanisms involved in HIV OB pathology in the context of HIV-1 infection.

Our finding of increased levels of T-cell activation is consistent with previous reports implicating peripheral T-cell activation with HIV-associated pulmonary dysfunction (20). Produced by activated T cells (24), soluble CD40-ligand binds CD40 expressing cells promoting inflammatory responses (25) alongside expression of adhesion molecules and MMPs (26). sCD25 (IL2-RA) correlates well with surface CD25 expression (27) and is a marker of activated T-regulatory cells (CD25+) (28). sCD25 has been used as a marker of disease severity in a number of inflammatory conditions (29) and is associated with an expanded Th17 response (30), which recruits pathogenetic T cells to sites of inflammation in inflammatory diseases such as asthma (31). In non-atopic asthma patients sCD25 is associated with FEV_1. These patients typically experience broncho-obstructive reactions to inflammatory stimuli similar to HCLD which normalises with sCD25 (32). Overall, our results suggest that treatment strategies which reduce the levels of T-cell activation in participants with HCLD may be beneficial.

Our findings suggest that HCLD is associated with elevated levels of several immune activation markers. CRP has previously been associated with pulmonary dysfunction in individuals with HIV infection (19,20) and is associated with FEV_1 decline at the population

level (33). Long term exposure to elevated levels of IP-10, which is likely in this population due to the strong association between IP-10 and active HIV replication, has been shown to cause bronchiolitis-like inflammation (34). In chronic obstructive pulmonary disease (COPD), tissue injury is thought to be promoted by IFN-γ through release of MMP from activated macrophages (35). Furthermore, elevated immune activation in HCLD is likely driven by HIV-associated increases in gut lumen permeability leading to microbial translocation. As a marker of monocyte activation in response to lipopolysaccharide, elevated sCD14 levels in the HCLD participants is suggestive of increased microbial translocation in individuals with CLD. sCD14 has previously been associated with airflow limitation and combined mosaic attenuation on chest computerised tomography (CT) scan, consistent with obliterative bronchiolitis (12). We describe an increased inflammatory state in individuals with HCLD and highlight that further studies assessing microbial translocation in this population are warranted.

In a similar study, Attia *et al* recently proposed that chronic inflammation may cause endothelial disruption that drives HCLD (12). Endothelial activation is well described in HIV-infected individuals (36), and is particularly marked in perinatal infection (37). E-selectin reflects the activation of endothelial cells and was elevated in HCLD participants in our study. In emphysema patients endothelial cells release pro-inflammatory cytokines such as TNF-alpha and IL-1-beta that contribute to CLD development (38). We also report elevations in D-Dimer, a fibrinogen breakdown product which has been associated with HIV all-cause mortality and acute execrations in patients with interstitial lung disease (39,40). D-Dimer is

261 also strongly correlated with endothelial dysfunction, microbial translocation and sCD14
262 (41–43).

263 Owing to the essential role of the pulmonary extracellular matrix (ECM) for normal lung
264 function, the association of several markers of extracellular matrix degradation (MMP1, 7 &
265 10) with HCLD is of interest (44). The concentration of MMPs from bronchoalveolar lavage
266 (BAL) fluid samples have been associated with radiological markers of small airway disease
267 and emphysema severity (45) and MMP-7 has been shown to promote pulmonary fibrosis
268 (46). MMP-10 is expressed by multiple cell types in response to infection (47), and likely
269 represents increased immune activation. MMP-1 is elevated in subjects with COPD and
270 children with pulmonary TB (45) but is responsive to TB treatment (48).

271

272 As soluble biomarkers do not act in isolation, we sought to study the relationships between
273 biomarkers in cases and controls to understand better the activated pathways that may
274 drive pathology. We suggest that biomarkers of high centrality in the HCLD group offer the
275 best potential for therapies aimed at reducing systemic immune activation. Furthermore,
276 principal component analysis confirms the dominance of biomarkers associated with T-cell
277 activation in this study.

278

279 There are several limitations to this study. The cross-sectional design means that the
280 direction of temporal relationships is unknown. Due to the exploratory nature of our study
281 we did not correct for multiple testing, and results must be interpreted cautiously. Ward *et*
282 *al* found no relationship between bronchoalveolar lavage (BAL) and blood sCD14 (29),

283 indicating that the levels of plasma soluble biomarkers may not represent local levels in
284 relevant organs. High-resolution computed tomographic scans and BAL sampling from
285 BREATHE participants would allow us better to describe the phenotype of HCLD and
286 describe local inflammation within the cohort. Despite this, our study furthers the
287 understanding of pathways associated with HCLD in this children and adolescents living with
288 PHIV, and suggests that studies aimed at characterising T-cell activation in this cohort may
289 be of benefit.

290

291 In conclusion, systemic inflammation, particularly T-cell activation, is associated with HCLD
292 in older children and adolescents on stable ART from SSA. Inflammation and immune
293 activation markers alongside markers of extracellular matrix degradation and cellular
294 adhesions are associated with reduced lung function. These results act as a first step to
295 identifying potential targets for therapeutic modalities that may be capable of preventing
296 the decline of lung function in this population, and highlight several probable pathways
297 associated with HCLD in this population.

298

299

300 **Acknowledgements**: This work was supported by Research Council of Norway, through their
301 GLOBVAC programme, with additional support from Helse Nord. RAF is funded by the
302 Wellcome Trust through a senior fellowship in clinical science. DHB is funded by the
303 Wellcome Trust PhD studentship in Genomic Medicine and Statistics

304

Contributions of authors; Initials (contribution). DHB (data collection, analysis and writing); ES (conception, writing and review); AMR, VS (data analysis, management and review); TF,TJG, LGN, TB, GMH, SRJ, RAF study conception, design, implementation and review; (LMY data collection, study design, review).

1. McHugh G, Rylance J, Mujuru H, Nathoo K, Chonzi P, Dauya E, et al. Chronic Morbidity among Older Children and Adolescents at Diagnosis of HIV Infection. J Acquir Immune Defic Syndr. 2016;73(3):275–81.

2. Ferrand RA, Miller RF, Kaski JP, Hakim J, Matenga J, Nathoo K, et al. Chronic Lung disease in adolescents with delayed diagnosis of vertically acquired HIV infection. Clin Infect Dis. 2013;56(4):576–82.

3. Desai SR, Nair A, Rylance J, Mujuru H, Nathoo K, Mchugh G, et al. Human Immunodeficiency Virus-Associated Chronic Lung Disease in Children and Adolescents in Zimbabwe : Chest Radiographic and High-Resolution Computed Tomographic Findings. Clin Infect Dis. 2018;66(2).

4. Majonga ED, Rehman AM, Simms V, Mchugh G, Mujuru HA, Nathoo K, et al. High prevalence of echocardiographic abnormalities in older HIV-infected children taking antiretroviral therapy. AIDS. 2018;0.

5. Ferrand RA, Munaiwa L, Matsekete J, Bandason T, Nathoo K, Ndhlovu CE, et al. Undiagnosed HIV Infection among Adolescents Seeking Primary Health Care in Zimbabwe. Clin Infect Dis [Internet]. 2010;51(7):844–51. Available from: https://academic.oup.com/cid/article-lookup/doi/10.1086/656361

6. Githinji LN, Gray DM, Hlengwa S, Myer L, Zar HJ. Lung function in South African adolescents infected Perinatally with HIV and treated long-term with antiretroviral

328 therapy. Ann Am Thorac Soc. 2017;14(5):722–9.

329 7. Frigati LJ, Brown K, Mahtab S, Githinji L, Gray D, Zühlke L, et al. Multisystem
330 impairment in South African adolescents with Perinatally acquired HIV on
331 antiretroviral therapy (ART). J Int AIDS Soc. 2019;22(8).

332 8. Githinji LN, Gray DM, Zar HJ. Lung function in HIV-infected children and
333 adolescents. Pneumonia. 2018;10(1):1–10.

334 9. Rylance J, McHugh G, Metcalfe J, Mujuru H, Nathoo K, Wilmore S, et al. Chronic
335 lung disease in HIV-infected children established on antiretroviral therapy. Aids.
336 2016;30(18):2795–803.

337 10. Triant VA, Meigs JB, Grinspoon SK. Association of C-reactive protein and HIV
338 infection with acute myocardial infarction. J Acquir Immune Defic Syndr.
339 2009;51(3):268–73.

340 11. Fitzpatrick ME, Nouraie M, Gingo MR, Camp D, Kessinger CJ, Sincebaugh JB, et al.
341 Novel relationships of markers of monocyte activation and endothelial dysfunction
342 with pulmonary dysfunction in HIV-infected persons. AIDS. 2016 Jun;30(9):1327–39.

343 12. Attia EF, Bhatraju PK, Triplette M, Kosamo S, Maleche-Obimbo E, West TE, et al.
344 Endothelial Activation, Innate Immune Activation, and Inflammation Are Associated
345 With Postbronchodilator Airflow Limitation and Obstruction Among Adolescents
346 Living With HIV. J Acquir Immune Defic Syndr. 2020;83(3):267–77.

347 13. Wang Y, Xu1 J, Meng1 Y, Adcock IM, Yao1 X. Role of inflammatory cells in airway
348 remodeling in COPD. 2018;3341–8.

349 14. Gonzalez-Martinez C, Kranzer K, McHugh G, Corbett EL, Mujuru H, Nicol MP, et al.
350 Azithromycin versus placebo for the treatment of HIV-associated chronic lung disease

in children and adolescents (BREATHE trial): Study protocol for a randomised controlled trial. Trials. 2017;18(1):1–8.

15. McHugh G, Rehman AM, Simms V, Gonzalez-Martinez C, Bandason T, Dauya E, et al. Chronic lung disease in children and adolescents with HIV: a case-control study. Trop Med Int Heal. 2020;00(00):1–10.

16. Ferrand RA, McHugh G, Rehman AM, Mujuru H, Simms V, Majonga ED, et al. Effect of Once-Weekly Azithromycin vs Placebo in Children With HIV-Associated Chronic Lung Disease: The BREATHE Randomized Clinical Trial. JAMA Netw open [Internet]. 2020 Dec 1;3(12):e2028484–e2028484. Available from: https://pubmed.ncbi.nlm.nih.gov/33331916

17. Cole TJ, Freeman J V., Preece MA. British 1990 growth reference centiles for weight, height, body mass index and head circumference fitted by maximum penalized likelihood. Stat Med. 1998;17(4):407–29.

18. Lê S, Josse J, Husson F. FactoMineR: An R Package for Multivariate Analysis. J Stat Software; Vol 1, Issue 1 [Internet]. 2008; Available from: https://www.jstatsoft.org/v025/i01

19. North CM, Muyanja D, Kakuhikire B, Tsai AC, Russell P, Hunt PW, et al. Systemic Inflammation, Immune Activation and Impaired Lung Function among People Living with HIV in Rural Uganda. 2019;78(March 2018):543–8.

20. Fitzpatrick ME, Singh V, Bertolet M, Lucht L, Kessinger C, Michel J, et al. Relationships of pulmonary function, inflammation, and T-cell activation and senescence in an HIV-infected cohort. AIDS. 2014 Nov;28(17):2505–15.

21. Colom AJ, Maffey A, Bournissen FG, Teper A. Pulmonary function of a paediatric

cohort of patients with postinfectious bronchiolitis obliterans. A long term follow-up. Thorax. 2015;70(2):169–74.

22. Rosewich M, Zissler UM, Kheiri T, Voss S, Eickmeier O, Schulze J, et al. Airway inflammation in children and adolescents with bronchiolitis obliterans. Cytokine [Internet]. 2015;73(1):156–62. Available from: http://dx.doi.org/10.1016/j.cyto.2014.10.026

23. Kelly K, Hertz MI. OBLITERATIVE BRONCHIOLITIS. Clin Chest Med [Internet]. 1997 Jun 1 [cited 2020 Jan 20];18(2):319–38. Available from: https://www.sciencedirect.com/science/article/abs/pii/S0272523105703828?via%3Dihub

24. Aukrust P, Müller F, Ueland T, Berget T, Aaser E, Brunsvig A, et al. Enhanced levels of soluble and membrane-bound CD40 ligand in patients with unstable angina: Possible reflection of T lymphocyte and platelet involvement in the pathogenesis of acute coronary syndromes. Circulation. 1999;100(6):614–20.

25. Chung HW, Lim JB. Clinical significance of elevated serum soluble CD40 ligand levels as a diagnostic and prognostic tumor marker for pancreatic ductal adenocarcinoma. J Transl Med. 2014;12(1):1–9.

26. Sousa AE, Chaves AF, Doroana M, Antunes F, Victorino RMM. Early reduction of the over-expression of CD40L, OX40 and Fas on T cells in HIV-1 infection during triple anti-retroviral therapy: Possible implications for lymphocyte traffic and functional recovery. Clin Exp Immunol. 1999;116(2):307–15.

27. Junghans RP, Waldmann TA. Metabolism of Tac (IL2Rα): Physiology of cell surface shedding and renal catabolism, and suppression of catabolism by antibody binding. J Exp Med. 1996;183(4):1587–602.

398 28. Corrigan CJ, Kay AB. CD4 T lymphocyte activation in acute severe asthma. Int Arch
399 Allergy Appl Immunol. 1991;94(1–4):270–1.

400 29. Ward C, Walters EH, Zheng L, Whitford H, Williams TJ, Snell GI. Increased soluble
401 CD14 in bronchoalveolar lavage fluid of stable lung transplant recipients. Eur Respir J.
402 2002;19(3):472–8.

403 30. Russell SE, Moore AC, Fallon PG, Walsh PT. Soluble IL-2Rα (sCD25) Exacerbates
404 Autoimmunity and Enhances the Development of Th17 Responses in Mice. PLoS One.
405 2012;7(10):1–9.

406 31. Tesmer LA, Lundy SK, Sarkar S, Fox DA. Th17 cells in human disease. Immunol
407 Rev. 2008;223(1):87–113.

408 32. Lasalle P, Sergant M, Delneste Y, Cosset P, Wallaert B, Zandecki M, et al. Levels of
409 soluble IL-2 receptor in plasma from asthmatics. Correlations with blood eosinophilia,
410 lung function, and corticosteroid therapy. Clin Exp Immunol. 1992;87(2):266–71.

411 33. Shaaban R, Kony S, Driss F, Leynaert B, Soussan D, Pin I, et al. Change in C-reactive
412 protein levels and FEV1 decline: A longitudinal population-based study. Respir Med.
413 2006;100(12):2112–20.

414 34. Agostini C, Calabrese F, Rea F, Facco M, Tosoni A, Loy M, et al. CXCR3 and its
415 ligand CXCL10 are expressed by inflammatory cells infiltrating lung allografts and
416 mediate chemotaxis of T cells at sites of rejection. Am J Pathol. 2001;158(5):1703–11.

417 35. Gadgil A, Duncan SR. Role of T-lymphocytes and pro-inflammatory mediators in the
418 pathogenesis of chronic obstructive pulmonary disease. Int J COPD. 2008;3(4):531–
419 41.

420 36. Calza L, Pocaterra D, Pavoni M, Colangeli V, Manfredi R, Verucchi G, et al. Plasma

levels of VCAM-1, ICAM-1, E-Selectin, and P-Selectin in 99 HIV-positive patients versus 51 HIV-negative healthy controls. Vol. 50, Journal of acquired immune deficiency syndromes (1999). United States; 2009. p. 430–2.

37. Dirajlal-Fargo S, Sattar A, Kulkarni M, Bowman E, Funderburg N, McComsey GA. HIV-positive youth who are perinatally infected have impaired endothelial function. AIDS. 2017 Sep;31(14):1917–24.

38. Polverino F, Celli BR, Owen CA. COPD as an endothelial disorder: endothelial injury linking lesions in the lungs and other organs? Pulm Circ. 2018;8(1).

39. Kuller LH, Tracy R, Belloso W, De Wit S, Drummond F, Lane HC, et al. Inflammatory and coagulation biomarkers and mortality in patients with HIV infection. PLoS Med. 2008;5(10):1496–508.

40. Ishikawa G, Acquah SO, Salvatore M, Padilla ML. Elevated serum D-dimer level is associated with an increased risk of acute exacerbation in interstitial lung disease. Respir Med [Internet]. 2017;128:78–84. Available from: http://dx.doi.org/10.1016/j.rmed.2017.05.009

41. Funderburg NT, Mayne E, Sieg SF, Asaad R, Jiang W, Kalinowska M, et al. Increased tissue factor expression on circulating monocytes in chronic HIV infection: Relationship to in vivo coagulation and immune activation. Blood. 2010;115(2):161–7.

42. Baker J, Quick H, Hullsiek KH, Tracy R, Duprez D, Henry K, et al. Interleukin-6 and d-dimer levels are associated with vascular dysfunction in patients with untreated HIV infection. HIV Med. 2010;11(9):608–9.

43. He W, Castiblanco J, Walter EA, Okulicz JF, Ahuja SK. Mendelian randomization: Potential use of genetics to enable causal inferences regarding HIV-associated

biomarkers and outcomes. Curr Opin HIV AIDS. 2010;5(6):545–59.

44. Burgstaller G, Oehrle B, Gerckens M, White ES, Schiller HB, Eickelberg O. The instructive extracellular matrix of the lung: Basic composition and alterations in chronic lung disease. Eur Respir J [Internet]. 2017;50(1). Available from: http://dx.doi.org/10.1183/13993003.01805-2016

45. Ostridge K, Williams N, Kim V, Bennett M, Harden S, Welch L, et al. Relationship between pulmonary matrix metalloproteinases and quantitative CT markers of small airways disease and emphysema in COPD. 2016;126–32.

46. Craig VJ, Zhang L, Hagood JS, Owen CA. Matrix metalloproteinases as therapeutic targets for idiopathic pulmonary fibrosis. Am J Respir Cell Mol Biol. 2015;53(5):585–600.

47. Kassim SY, Gharib SA, Mecham BH, Birkland TP, Parks WC, McGuire JK. Individual matrix metalloproteinases control distinct transcriptional responses in airway epithelial cells infected with Pseudomonas aeruginosa. Infect Immun. 2007;75(12):5640–50.

48. Kumar NP, Moideen K, Viswanathan V, Shruthi BS, Sivakumar S, Menon PA, et al. Elevated levels of matrix metalloproteinases reflect severity and extent of disease in tuberculosis-diabetes co-morbidity and are predominantly reversed following standard anti-tuberculosis or metformin treatment. BMC Infect Dis. 2018;18(1):1–10.

Figures and Tables

Table 1: List of soluble biomarkers studied

Pathway	Biomarker	Abbreviation
Immune Activation	Beta-2-Microglobulin	B2M
	Granulocyte Colony Stimulating Factor	GCSF
	Interferon induced protein-10	IP-10
	C-Reactive Protein	CRP
	Interferon Gamma	INF-G
Monocyte Activation	Soluble CD14	sCD14
	Soluble CD163	sCD163
T-Cell Activation	Soluble CD40-Ligand	sCD40-Ligand
	Soluble CD27	sCD27
	Soluble CD25	sCD25
	Soluble CCL5	sCCL5
Endothelial Activation & Cellular Adhesion	E-Selectin	E-Selectin
	P-Selectin	P-Selectin
	Vascular cell adhesion molecule 1	VCAM-1
	Intracellular cell adhesion molecule 1	ICAM-1
	Vascular endothelial growth factor	VEGF
Coagulation	D-Dimer and fibrin degradation products	D-Dimer +
Apoptosis	Fas	Fas
	Growth differentiation factor 15	GDF-15
Angiogenesis	Angiopoietin-1	ANG-1
Extracellular matrix	Matrix metalloproteinase-1	MMP-1
	Matrix metalloproteinase-3	MMP-3

	Matrix metalloproteinase-7	MMP-7
	Matrix metalloproteinase-8	MMP-8
	Matrix metalloproteinase-10	MMP-10
	Matrix metalloproteinase-12	MMP-12

466

467

468

469

470 Table 2: Clinical, demographic and anthropometric characteristics of participants included

471 from the BREATHE trial.

	Controls (n=74)	**Cases (n=336)**	**P Value**
Age at Enrolment – Mean (SD)	14.9 (3.6)	15.0 (3.2)	0.833
Sex FEMALE, N (%)	46 (62.2)	166 (49.4)	0.063
Zimbabwe, N (%)	55 (74.3)	241 (71.7)	0.758
Malawi, N (%)	19 (25.7)	95 (28.3)	
Height for Age Z-Score (Median (IQR)	-1.5 (1.2)	-2.0 (1.4)	<0.001
Stunted, N (%)	22 (29.7)	168 (50.0)	0.002
Weight for Age Z-Score (Median (IQR)	-1.1 (1.2)	-2.2 (1.5)	<0.001
Wasting, N (%)	14 (18.9)	176 (52.4)	<0.001

CD4 T Cell Count (Cells /mm^3) > 350, N (%) *	38 (51.4)	145 (43.2)	0.248
HIV Viral Control (<200 copies/ml), N (%)*	63 (85.1)	297 (88.4)	0.563
Log-10 HIV Viral Load Copies/ml, Median (IQR)*	2.0 (1.8)	2.6 (2.5)	0.087
FEV_1 Z-Score, Median (IQR)	0.5 (0.6)	-2.0 (1.0)	<0.001
FEV1/FVC z-score (Mean (SD)	0.2 (0.8)	-0.7 (1.1)	<0.001
FEV % Predicted Mean (SD)	107.7 (6.1)	72.8 (9.9)	<0.001
Duration ART in Years, Median (IQR)	6.8 (5.0)	6.4 (3.2)	0.758
Ever Treated for TB, N (%)*	9 (12.2)	97 (28.9)	0.005
ART Regime, ATV/LPV/PI N (%)	8 (10.8)	88 (26.2)	0.008
ART Regime EFV/NVP N (%)*	65 (87.8)	247 (73.5)	

IQR = Interquartile Range, N= Number, ATV = Atanavir, LPV=Lopanavir, PI= Protease inhibitor EFV= Efavirenz NVP = Nevirapine, * = contains missing data. One data point missing for CD4, Ever Treated for TB and ART regime for both groups. Two data points missing for HIV viral load (imputed using mean imputation).

479

480

481

482

483 Table 3. Biomarkers significantly associated with HCLD in the case control analysis

484

	Univariate Logistic Regression, OR (CI, P)	Adjusted Logistic Regression, OR (CI, P)
MMP-1	1.41 (1.09-1.84, p= .009)	1.36 (1.04-1.79, p= .028)
MMP-7	1.52 (1.18-1.98, p=0= .001)	1.42 (1.07-1.90, p= .015)
MMP-10	1.70 (1.29-2.25, p< .001)	1.61 (1.20-2.19, p= .002)
Angiopoietin-1	1.49 (1.19-1.90, p= .001)	1.53 (1.20-1.96, p= .001)
sCCL5	1.36 (1.06-1.75, p= .015)	1.37 (1.05-1.80, p= .023)
sCD14	1.96 (1.51-2.57, p< .001)	2.23 (1.66-3.05, p< .001)
sCD25	2.74 (2.01-3.81, p< .001)	2.85 (2.00-4.19, p< .001)
sCD27	2.26 (1.65-3.16, p< .001)	2.05 (1.48-2.91, p< .001)
sCD40-Ligand	2.89 (2.12-4.06, p< .001)	2.96 (2.12-4.25, p< .001)
CRP	1.55 (1.20-2.04, p= .001)	1.48 (1.12-1.98, p= .006)
IP-10/CXCL10	1.93 (1.42-2.69, p< .001)	1.89 (1.36-2.72, p< .001)

D-Dimer +	1.68 (1.27-2.25, p< .001)	1.68 (1.25-2.29, p= .001)
E-Selectin	2.08 (1.57-2.81, p< .001)	2.05 (1.52-2.82, p< .001)
FAS	1.59 (1.23-2.08, p= .001)	1.59 (1.21-2.12, p= .001)
GCSF	0.75 (0.58-0.97, p= .032)	0.68 (0.50-0.91, p= .010)
IFN-γ	1.75 (1.35-2.28, p< .001)	2.63 (1.77-4.12, p< .001)
VCAM-1	1.70 (1.29-2.30, p< .001)	1.56 (1.18-2.10, p= .003)
VEGF	0.79 (0.61-1.02, p= .070)	0.73 (0.55-0.95, p= .022)

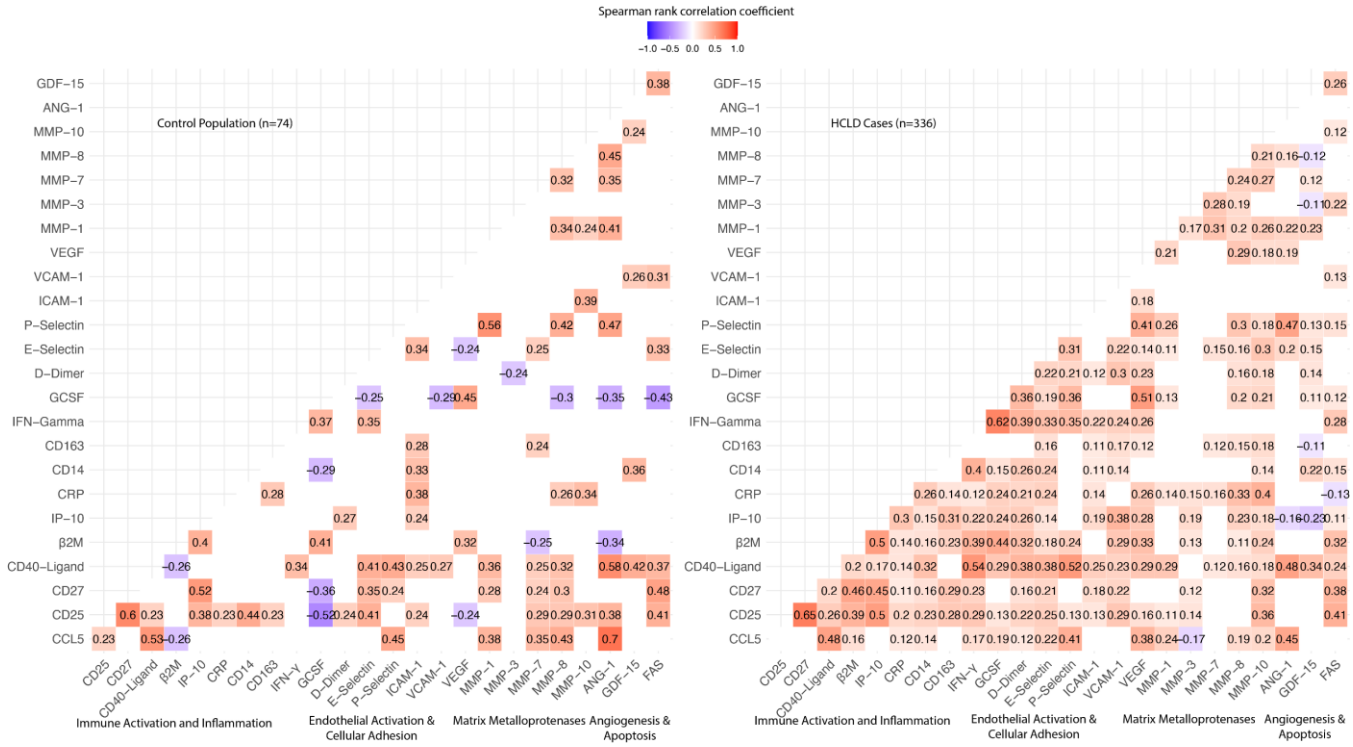
Table 3. Univariate and adjusted logistic regression of scaled $\log_{10}$ Biomarker and HCLD status. Odds ratios represent change in odds per one standard deviation increase in biomarker * adjusted models include Age, Sex, Study site, Height for age z-scores, HIV viral suppression and having ever been treated for TB. Only biomarkers with p< .05 in adjusted analysis are shown.

Table 4: Biomarkers associated with FEV1 z-score in participants with HCLD

	Univariate β ± SE	Univariate P-Value	Adjusted β ± SE	Adjusted P-Value
MMP-8	-0.122 ± 0.039	0.002	-0.097 ± 0.039	0.012
MMP-10	-0.161 ± 0.038	0	-0.132 ± 0.04	0.001
ANG-1	0.08 ± 0.039	0.041	0.077 ± 0.039	0.046
CRP	-0.149 ± 0.038	0	-0.128 ± 0.038	0.001
IP-10	-0.089 ± 0.039	0.023	-0.08 ± 0.04	0.048
E-Selectin	-0.104 ± 0.039	0.008	-0.082 ± 0.039	0.037
Fas	0.083 ± 0.039	0.035	0.082 ± 0.039	0.033
GCSF	-0.106 ± 0.039	0.007	-0.11 ± 0.039	0.006
VCAM-1	-0.105 ± 0.039	0.007	-0.09 ± 0.039	0.021
VEGF	-0.117 ± 0.039	0.003	-0.113 ± 0.039	0.004

Multivariate Linear Regression of normalised and scaled biomarkers and FEV_1 z-score in the HCLD group. Covariates included age, sex, having ever been treated for TB, study site, height for age z-score and supressed viral load. Only biomarkers with statistically significant associations ($p < .05$) are shown.

Figure 1: Spearman rank correlations between different biomarkers studied split by case control status



A) Control group and B) HCLD group. Number within square indicates spearman rank correlation coefficient between biomarkers. Non-significant biomarker pairs ($p > .05$) are represented by blank squares. MMP-12 dropped from analysis due to high number of observations below the limit of detection.

544

545

546

547

548

549 **Supplementary Figures/Tables:**

550

551 Supplementary Table 1: Untransformed biomarker levels (pg/ml) in cases and controls

Biomarker	Controls (pg/ml), Median (IQR) (n=74)	Cases (pg/ml), Median (IQR) (n=336)	P value
MMP-1	1232.7 (1139.6)	1577.7 (1811.6)	0.021
MMP-3	3570.5 (3392.8)	3506.9 (3865.6)	0.929
MMP-7	1173.8 (685.0)	1490.0 (939.7)	<0.001
MMP-8	2868.5 (5383.5)	4013.2 (6760.7)	0.049
MMP-10	736.7 (607.9)	1071.4 (844.9)	<0.001
MMP-12	21.6 (19.8)	28.5 (19.8)	0.002
Angiopoietin-1	11669.8 (12856.8)	14868.1 (12936.1)	0.013
Beta-2-Microglobulin	2577312.1 (1026666.1)	2708684.2 (1890922.2)	0.166
CCL5	36568.1 (32328.7)	44381.6 (26809.5)	0.022
sCD14	1630166.3 (677911.7)	1959313.6 (863922.8)	<0.001
sCD25	423.1 (272.4)	648.7 (367.6)	<0.001
sCD27	6318.9 (3315.5)	8367.9 (4744.5)	<0.001

sCD40-Ligand	2912.1 (1184.9)	4454.5 (2037.8)	<0.001
sCD163	704196.9 (557007.0)	824710.1 (644962.8)	0.014
CRP	351638.7 (810154.9)	681709.5 (3846846.5)	0.001
IP-10	77.8 (57.5)	115.9 (125.6)	<0.001
D-Dimer	1306651.5 (1021554.2)	1903081.0 (2002126.2)	<0.001
E-Selectin	28610.0 (13313.1)	35903.1 (18481.5)	<0.001
FAS	4517.6 (1440.8)	5094.7 (1870.0)	0.001
GCSF	104.0 (66.8)	81.3 (60.6)	0.016
GDF-15	797.3 (877.2)	828.2 (1125.3)	0.864
ICAM-1	247118.8 (231948.2)	293854.2 (287375.8)	0.139
IFN-γ	77.8 (38.4)	108.4 (77.5)	<0.001
P-Selectin	28284.5 (12424.7)	30784.6 (13095.8)	0.011
VCAM-1	830248.8 (553079.4)	1065117.6 (698897.2)	<0.001
VEGF	69.3 (50.6)	49.3 (40.0)	0.027

552 Differences between group assessed with Kruskal Wallis non-parametric test

553

554 Supplementary Table 2: Tests of association between biomarker levels and HCLD status

Biomarker (Log_{10}, Scaled)	Univariate Logistic Regression, OR (CI, P)	Adjusted Logistic Regression, OR (CI, P)
MMP-1	1.41 (1.09-1.84, p=0.009)	1.36 (1.04-1.79, p=0.028)
MMP-3	1.02 (0.79-1.31, p=0.902)	1.05 (0.78-1.44, p=0.754)
MMP-7	1.52 (1.18-1.98, p=0.001)	1.42 (1.07-1.90, p=0.015)
MMP-8	1.24 (0.96-1.61, p=0.095)	1.14 (0.88-1.48, p=0.340)
MMP-10	1.70 (1.29-2.25, p<0.001)	1.61 (1.20-2.19, p=0.002)
MMP-12	1.25 (0.98-1.58, p=0.063)	1.25 (0.95-1.63, p=0.101)
Angiopoietin-1	1.49 (1.19-1.90, p=0.001)	1.53 (1.20-1.96, p=0.001)
Beta-2-Microglobulin	1.14 (0.89-1.46, p=0.302)	1.03 (0.75-1.42, p=0.837)
sCCL5	1.36 (1.06-1.75, p=0.015)	1.37 (1.05-1.80, p=0.023)
sCD14	1.96 (1.51-2.57, p<0.001)	2.23 (1.66-3.05, p<0.001)
sCD25	2.74 (2.01-3.81, p<0.001)	2.85 (2.00-4.19, p<0.001)
sCD27	2.26 (1.65-3.16, p<0.001)	2.05 (1.48-2.91, p<0.001)
sCD40-Ligand	2.89 (2.12-4.06, p<0.001)	2.96 (2.12-4.25, p<0.001)
sCD163	1.33 (1.04-1.72, p=0.024)	1.32 (1.00-1.76, p=0.052)
CRP	1.55 (1.20-2.04, p=0.001)	1.48 (1.12-1.98, p=0.006)

IP-10/CXCL10	1.93 (1.42-2.69, p<0.001)	1.89 (1.36-2.72, p<0.001)
D-Dimer +	1.68 (1.27-2.25, p<0.001)	1.68 (1.25-2.29, p=0.001)
E-Selectin	2.08 (1.57-2.81, p<0.001)	2.05 (1.52-2.82, p<0.001)
FAS	1.59 (1.23-2.08, p=0.001)	1.59 (1.21-2.12, p=0.001)
GCSF	0.75 (0.58-0.97, p=0.032)	0.68 (0.50-0.91, p=0.010)
GDF-15	1.04 (0.81-1.34, p=0.778)	0.99 (0.75-1.32, p=0.950)
ICAM-1	1.05 (0.81-1.33, p=0.701)	1.00 (0.75-1.30, p=0.981)
IFN-γ	1.75 (1.35-2.28, p<0.001)	2.63 (1.77-4.12, p<0.001)
P-Selectin	1.29 (0.99-1.69, p=0.064)	1.22 (0.92-1.63, p=0.175)
VCAM-1	1.70 (1.29-2.30, p<0.001)	1.56 (1.18-2.10, p=0.003)
VEGF	0.79 (0.61-1.02, p=0.070)	0.73 (0.55-0.95, p=0.022)

555 Logistic regression results for all biomarkers included in the study. Abbreviations: OR=Odds
556 Ratio, CI=Confidence Interval, P = P-Value). Adjusted analysis controlled for age, sex, study
557 site, height for age z-scores, HIV viral suppression and having ever been treated for TB.

558

559

560

561

562

563

564

565 Supplementary Table 3: Association of biomarkers with FEV_1 z-score

Biomarker	Univariate Linear Regression ± SE	P-Value	Adjusted Linear regression ± SE	P-value
MMP-1	-0.063 ± 0.039	0.109	-0.034 ± 0.039	0.374
MMP-3	0.021 ± 0.039	0.595	0.013 ± 0.042	0.767
MMP-7	-0.059 ± 0.039	0.132	-0.021 ± 0.039	0.586
MMP-8	-0.122 ± 0.039	0.002	-0.097 ± 0.039	0.012
MMP-10	-0.161 ± 0.038	0	-0.132 ± 0.04	0.001
MMP-12	-0.022 ± 0.039	0.571	-0.026 ± 0.039	0.502
Angiopoietin-1	0.08 ± 0.039	0.041	0.077 ± 0.039	0.046
Beta-2-Microglobulin	-0.055 ± 0.039	0.161	-0.063 ± 0.044	0.151
sCCL5	-0.04 ± 0.039	0.311	-0.04 ± 0.039	0.304
sCD14	0.046 ± 0.039	0.238	0.05 ± 0.039	0.2
sCD25	-0.07 ± 0.039	0.073	-0.035 ± 0.041	0.39
sCD27	-0.071 ± 0.039	0.069	-0.04 ± 0.04	0.311
sCD40-Ligand	-0.034 ± 0.039	0.389	-0.026 ± 0.039	0.508
sCD163	0.018 ± 0.039	0.64	0.023 ± 0.04	0.563
CRP	-0.149 ± 0.038	0	-0.128 ± 0.038	0.001

IP-10/CXCL10	-0.089 ± 0.039	0.023	-0.08 ± 0.04	0.048
D-Dimer +	-0.062 ± 0.039	0.113	-0.055 ± 0.039	0.165
E-Selectin	-0.104 ± 0.039	0.008	-0.082 ± 0.039	0.037
FAS	0.083 ± 0.039	0.035	0.082 ± 0.039	0.033
GCSF	-0.106 ± 0.039	0.007	-0.11 ± 0.039	0.006
GDF-15	-0.028 ± 0.039	0.474	-0.009 ± 0.04	0.821
ICAM-1	-0.04 ± 0.039	0.312	-0.016 ± 0.039	0.679
IFN-γ	0.013 ± 0.039	0.735	-0.005 ± 0.042	0.905
P-Selectin	-0.028 ± 0.039	0.473	-0.023 ± 0.039	0.55
VCAM-1	-0.105 ± 0.039	0.007	-0.09 ± 0.039	0.021
VEGF	-0.117 ± 0.039	0.003	-0.113 ± 0.039	0.004

566

567 Univariate and adjusted linear regression results for biomarkers and FEV_1 z-score in the
568 HCLD group. SE= Standard Error. Adjusted analysis controlled for age, sex, study site,
569 height for age z-scores, HIV viral suppression and having ever been treated for TB.

570

571

572

573

574

575

Supplementary Table 4 – PCA Results

	Whole Population (n=410)			HCLD Population (n=336)		
	Eigenvalue	% Variance	Cumulative Variance (%)	Eigenvalue	% Variance	Cumulative % Variance
Principal Component 1	5.155	20.622	20.622	5.075	20.3	20.3
Principal Component 2	2.61	10.442	31.064	2.758	11.03	31.331
Principal Component 3	2.158	8.631	39.695	2.02	8.081	39.412
Principal Component 4	1.85	7.402	47.096	1.672	6.686	46.098
Principal Component 5	1.325	5.302	52.398	1.355	5.422	51.52
Principal Component 6	1.29	5.16	57.558	1.282	5.127	56.647
Principal Component 7	1.037	4.15	61.707	1.05	4.199	60.846
Principal Component 8	0.991	3.962	65.67	1.019	4.076	64.922
Principal Component 9	0.894	3.577	69.246	0.953	3.813	68.734
Principal Component 10	5.155	3.291	72.537	0.858	3.431	72.165

Eigenvalues and percentage of variance explained for each factor derived from principal component analysis in whole population and HCLD population. Grey squares represent dimensions with eigenvalues <1 (not included in downstream analysis).

589 Supplementary Table 5: Results of ROC analysis showing sensitivity and specificity of
590 biomarkers measured for HCLD.

	Threshold	Sensitivity	Specificity	AUC
MMP-1	3.297	0.77	0.387	0.586
MMP-3	3.613	0.473	0.595	0.497
MMP-7	3.196	0.784	0.452	0.635
MMP-8	3.25	0.405	0.753	0.573
MMP-10	2.919	0.622	0.658	0.643
MMP-12	1.45	0.77	0.509	0.617
Angiopoietin-1	3.808	0.284	0.908	0.593
Beta-2-Microglobulin	6.505	0.797	0.396	0.551
sCCL5	4.509	0.486	0.747	0.585
sCD14	6.324	0.851	0.435	0.685
sCD25	2.651	0.595	0.812	0.74
sCD27	3.837	0.608	0.705	0.688
sCD40-Ligand	3.526	0.716	0.812	0.768
sCD163	5.941	0.689	0.485	0.591
CRP	6.143	0.824	0.396	0.626
IP-10/CXCL10	1.918	0.568	0.729	0.665
D-Dimer +	6.165	0.635	0.643	0.648
E-Selectin	4.493	0.622	0.679	0.687
FAS	3.715	0.757	0.47	0.627
GCSF	2.004	0.541	0.673	0.589

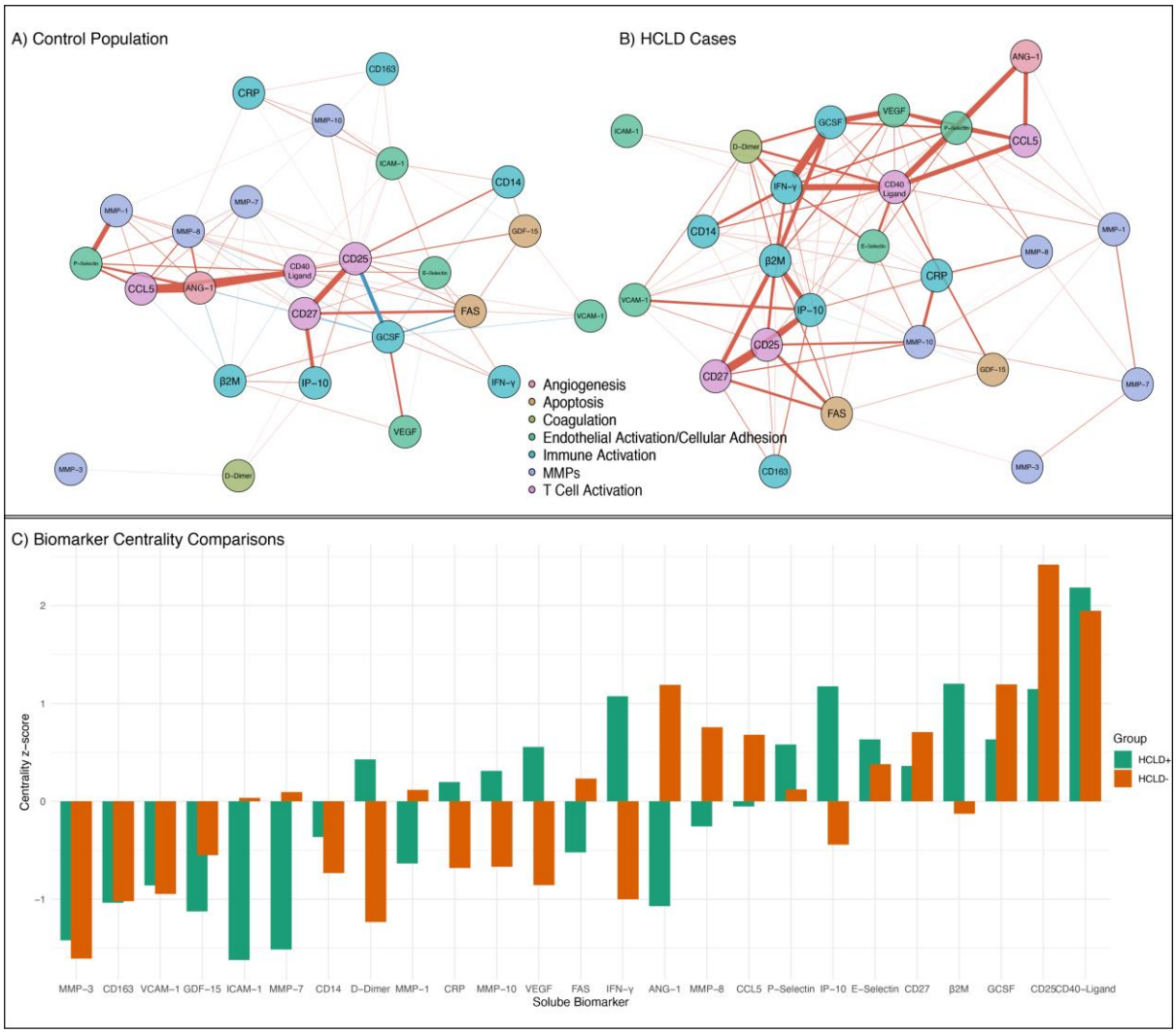
GDF-15	3.041	0.716	0.372	0.506
ICAM-1	5.593	0.797	0.339	0.555
IFN-γ	2.03	0.838	0.512	0.697
P-Selectin	4.47	0.635	0.562	0.595
VCAM-1	5.867	0.432	0.815	0.641
VEGF	1.824	0.554	0.682	0.582
Principal Component 1	-0.485	0.797	0.643	0.76
Principal Component 2	-0.405	0.716	0.387	0.541
Principal Component 3	-1.035	0.473	0.821	0.637
Principal Component 4	0.358	0.649	0.622	0.642
Principal Component 5	-0.554	0.838	0.345	0.589
Principal Component 6	0.192	0.568	0.604	0.578
Principal Component 7	0.219	0.716	0.452	0.58

Sensitivity and specificity refer to values at the specified threshold. Biomarker levels are $\log_{10}$ transformed.

597 Supplementary Figure 1: Network and centrality of biomarkers

598

599



600

601

602

603 Supplementary Figure 1. Network plots showing strength and direction of correlations
604 between biomarkers in A) Controls and B) Cases. The colour saturation and the width of the
605 edges corresponds to the absolute weight and scale relative to the strongest weight in the
606 graph. Nodes arranged by spring format. Biomarkers coloured by *a priori* biological

607 pathway. Only significant ($p < 0.05$) correlations greater than 0.2 are shown C) Standardized
608 (z-scores) centrality indices node strength for the networks in A/B. This indicates the
609 interconnectedness of each biomarker within the network.

610

611

612

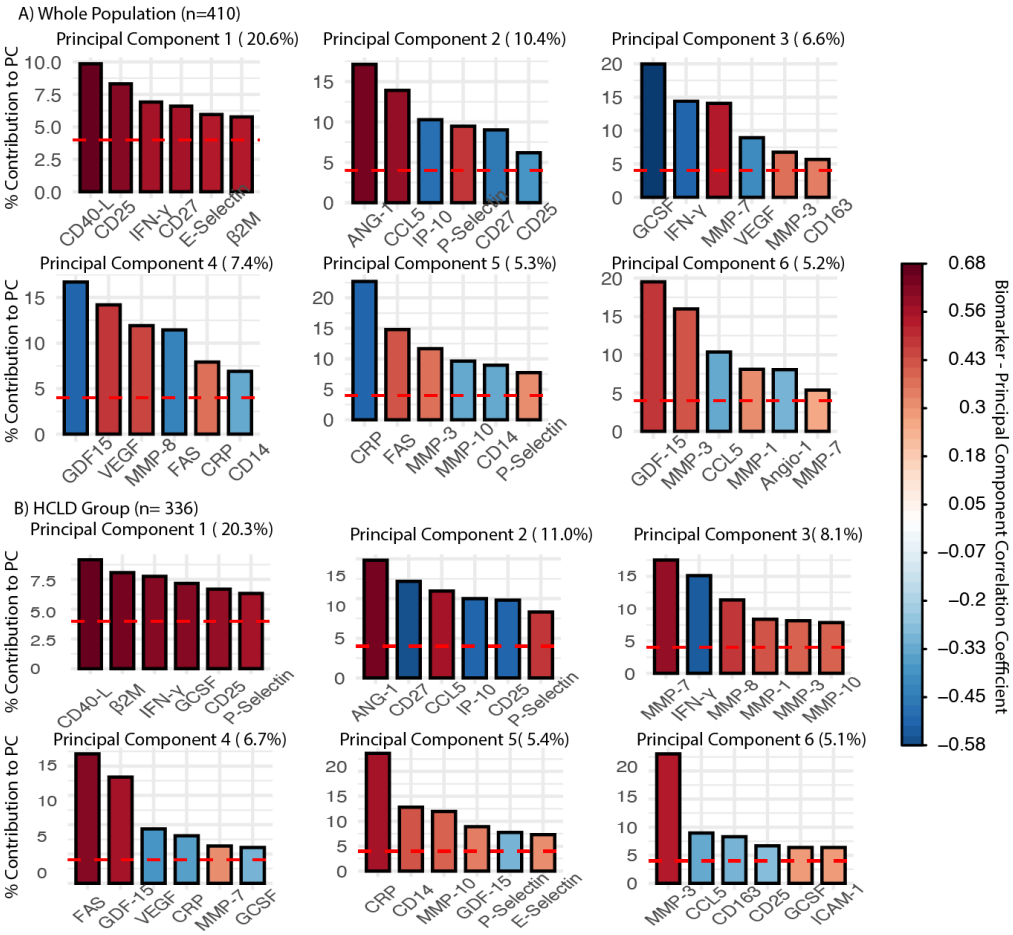
613

614

615

616 Supplementary Figure 2: Contribution of biomarkers to top six principle components in A)
617 Whole population and B) Participants with HCLD.

618



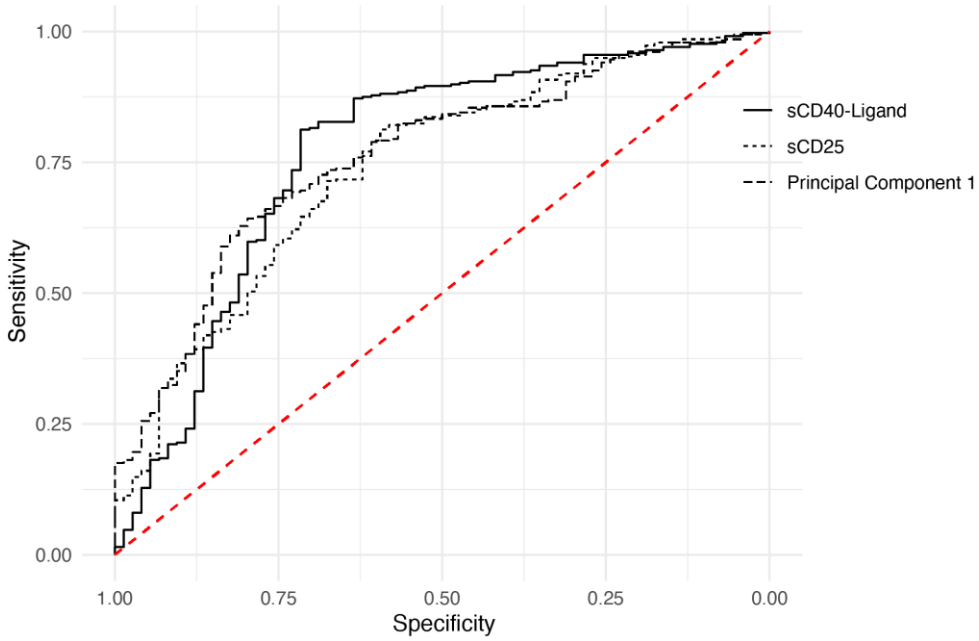
Percentages represent how much of variation of group biomarker data is explained by each principal component. Top 6 biomarkers contributing to each component shown. Dashed red line indicates contribution of biomarker if all equally contributing.

Supplementary Table 6: PCA components associated with HCLD and FEV-1 z-score in whole population and HCLD group respectively.

	Adjusted Logistic Regression Results (OR (95% CI, p))	Adjusted Linear Regression Results (β (SE), p)
Principal Component 1	1.54 (1.33-1.80, p<0.001)	-0.044 ± 0.018, p=0.014
Principal Component 2	0.97 (0.81-1.16, p=0.728)	0.004 ± 0.024, p=0.873
Principal Component 3	1.44 (1.16-1.81, p=0.001)	-0.042 ± 0.029, p=0.138
Principal Component 4	0.62 (0.49-0.77, p<0.001)	0.096 ± 0.03, p=0.001
Principal Component 5	0.71 (0.54-0.92, p=0.012)	-0.092 ± 0.036, p=0.011
Principal Component 6	0.74 (0.58-0.94, p=0.014)	-0.031 ± 0.036, p=0.395
Principal Component 7	1.47 (1.11-1.95, p=0.007)	0.082 ± 0.038, p=0.031
Principal Component 8		0.004 ± 0.039, p=0.911

All analyses include age, sex, site, height for age z score, HIV viral control and having ever been treated for TB as confounders.

Supplementary Figure 3: Receiver operating characteristics for biomarkers and exploratory PCA dimensions with AUC >0.7



ROC curve for top three performing variables. Diagonal line represents no predictive ability.