Full Title (120 Characters): Soluble Biomarkers associated with chronic lung disease in older children and adolescents with perinatal HIV infection

Running title (40 Characters): Soluble biomarkers associated with HCLD

Abstract Word Count: 230

Total Manuscript Word Count: 2701

Authors:

Dan HAMEIRI-BOWEN[1], MSc

Evgeniya SOVERSHAEVA[2,3], MD PhD

Trond FLAEGSTAD [2,3], PhD

Tore JARL GUTTEBERG [2,3], PhD

Lucky Gift NGWIRA[7], MPH

Victoria SIMMS[5], PhD

Andrea M REHMAN[5], PhD

Grace MCHUGH[4,6], MD

Tsitsi BANDASON[6], MSc

Rashida Abbas FERRAND [4,6], MD PhD

Sarah ROWLAND-JONES[1], DM

Louis-Marie YINDOM[1], PhD*

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, Department of Clinical Research, London, United Kingdom
5 London School of Hygiene and Tropical Medicine, International Statistics and Epidemiology Group, London, United Kingdom
6 Biomedical Research and Training Institute, Harare, Zimbabwe
7 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 Helse Nord. RAF is funded by the Wellcome Trust through a Senior Fellowship in Clinical Science. DHB is funded by the Wellcome Trust PhD Programme in Genomic Medicine and Statistics, grant number 108861/B/15/Z. AMR & VS are additionally funded by the UK Medical Research Council (MRC) and the UK department for international Development (DFID) under the MRC/DFID Concordant agreement, also part of the EDCTP2 programme supported by the European Union, Grant Ref: MR/R010161/1.

Correspondence:

Louis-Marie Yindom

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

Abstract (230/250 Words)

***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 twenty-six plasma soluble biomarkers and HCLD.

***Design**:* Case-control analysis of baseline biomarker data from 336 children and adolescents (6-19 years old) with perinatal HIV infection (PHIV) and HCLD (cases) and 74 age and sex-matched controls with PHIV but no CLD. HCLD was defined as having a forced expiratory volume in one second (FEV1) z-score <-1 with no reversibility.

***Methods**: Cryopreserved plasma collected at recruitment was used in a multiplex bead assay (Luminex) to measure baseline levels of soluble biomarkers. Logistic regression alongside data-reduction and techniques quantifying the interconnectedness of biomarkers were used to identify biomarkers associated with odds of HCLD.*

***Results**:* Biomarkers of general immune activation and inflammation (β2M, CRP, sCCL5, GCSF, IFN-γ, IP-10), T-Cell activation (sCD25, sCD27), platelet activation (sCD40-L), 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 and assessment of biomarker interconnectedness identified T-cell and platelet activation as centrally important to this association.

Conclusions: HCLD was associated with a large number of soluble biomarkers representing a range of different pathways. Our findings suggest a prominent role for T-cell and platelet activation in HCLD.

Introduction

The widespread use of combination antiretroviral therapy (ART) has led to a growing number of children 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 has reduced the incidence of pulmonary infections, there remains a substantial burden of chronic respiratory symptoms among 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 covering a wide spectrum of pathogenetic pathways with HCLD in children in Malawi and Zimbabwe. We also assessed how the associations between biomarkers changed in HCLD participants to better understand pathways dysregulated in disease. The predictive ability of individual biomarkers for HCLD was also assessed.

Methods:

This study was nested within the BREATHE trial (Bronchopulmonary function in response to azithromycin treatment for chronic lung disease in HIV-infected children) (ClinicalTrials.gov, NCT02426112), that investigated the impact of azithromycin therapy on lung function in children with HCLD. Full details of the trial 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 (Forced expiratory volume in one second (FEV1) 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 (TB) or potentially fatal conditions at time of screening were excluded. TB was screened using the Xpert MTB/RIF assay (Cepheid). Trial participants served as cases for the current study. Controls without HCLD (FEV1 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

Soluble biomarkers were measured from cryopreserved plasma stored at -80^{o}C. The full list of biomarkers measured and their abbreviations are reported in Supplementary 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). Briefly, plasma samples collected from heparinised blood were thawed on their first use, diluted appropriately, and the level of biomarkers assessed immediately in duplicates on a single MagPix instrument. Samples with measurement values falling outside the standard curve were repeated at appropriate dilutions. Those with consistently low levels of detection upon repeating were considered undetectable, and assigned half the minimum value measured for the specific biomarker under investigation.

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 standard deviations.

Spearman rank correlation coefficients between all soluble biomarkers were calculated in all participants and separately within cases. 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 node strength centrality (interconnectedness) of each biomarker was calculated from the absolute value of edge weights 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 FEV1 z-score as a continuous measure was assessed using linear regression. Variables associated with HCLD, FEV1 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-scores were included as continuous variables. Adjusted and unadjusted odds ratios, 95% confidence intervals (CI), regression coefficients and their standard errors are presented. Where appropriate, missing data in clinical covariates were imputed by mean imputation. Due to the exploratory nature of the study we did not correct for multiple testing.

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 (PC) were extracted and included in the logistic and linear

regression analyses 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 whole population 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= <.002), have ever been treated for TB (29.9% vs 12.2% p = .005) and to be on second line ART (25.9% vs 10.8%, p= .009) (Table 1). There was no difference in the proportion of individuals with suppressed HIV viral load between cases and controls (43.2% vs 51.4%, p=.248). A total of 59 (0.6%) biomarker measurements fell below the limit of detection. One measurement was missing. MMP-12 was dropped from data reduction owing to a reasonable number of values falling below the limit of detection (13%). Biomarker levels by group are presented in Figure 1A. Untransformed levels of each soluble biomarker (pg/ml) are reported in Supplementary Table 1). The centrality z-score for each soluble biomarker by group is presented in Figure 1D. sCD40-L was the most interconnected biomarker in cases and second most in controls (Figure 1B/C). Several

biomarkers showed noticeable differences between cases and controls (ANG-1, β2M, D-Dimer, IFN-γ, IP-10).

There was a high degree of correlation between biomarkers (Supplementary Figure 1). Separate exploratory PCA in all participants and just those with HCLD identified 7 and 8 principal components with eigenvalues >1 explaining 61 and 64% of the variation respectively (Supplementary Figure 2, Supplementary Table 2). These were extracted and used in downstream analysis.

18 soluble biomarkers and 6 principal components were associated with HCLD status. MMP-1, MMP-7, MMP-10, ANG-1, sCCL5, sCD14, sCD25, sCD27, sCD40-L, CRP, IP-10, D-Dimer, E-Selectin, Fas, IFN-γ, VCAM-1, PC1, PC3 and PC7 were associated with increased odds of HCLD. sCD40-L was associated with the largest increase in odds (OR=2.96 (95% CI= 2.21-4.25, p< .001)). GCSF, VEGF and PC4, PC5 and PC6 were associated with reduced odds of HCLD, GCSF was associated with the largest reduction (OR= 0.68 (95% CI= 0.50-0.91, p=0.010)) (Table 2 & Supplementary Table 3 & 5).

Among cases, MMP-8, MMP-10, ANG-1, CRP, IP-10, E-Selectin, Fas, GCSF, VCAM-1, VEGF, PC1 and PC5 were associated with reduced FEV1 z-score (Table 3 and Supplementary Table 4/5). The coefficients were largest for MMP-10 and CRP (β = -0.132 ± 0.04 and β = -0.128 ± 0.38, respectively). An increase of one standard deviation in Fas was associated with a small increase in FEV1 z-score (β = 0.082 ± 0.039, p= .028). Of these biomarkers, only MMP-8 was not associated with HCLD in the previous analyses.

Receiver operator characteristics (ROC) identified CD40-L, sCD25 and PCA principal component 1 as having area under the curve (AUC) greater than 0.7 (Supplementary Figure 3

& Supplementary Table 6). The best performing biomarker was log transformed sCD40-L 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 studies in children are consistent with constrictive obliterative bronchiolitis (OB) as the predominant underlying cause of HCLD (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 OB pathology in the context of HIV-1 infection further to those previously described (15,24).

Platelets are responsible for the vast majority of sCD40-L found in plasma (25) and play a key role in initiation and propagation inflammatory lung diseases (26). Excessive ligation of surface CD40 by sCD40-L drives an inflammatory endothelial cell phenotype (for which VCAM-1 is a marker of (26)) and increases expression of MMPs (27). sCD40-L is associated with a range of inflammatory conditions in the context of HIV (28,29) and has been postulated as a therapeutic target for HIV associated neuroinflammation (30). The high centrality of sCD40-L in both cases and controls is unsurprising considering the co-stimulatory properties of sCD40-L in both innate and adaptive immunity. The high sensitivity and specificity of sCD40-L for HCLD further underscores its central role and potential use as a biomarker for HCLD. Combined, our results highlight the importance of sCD40-L and/or cells involved in its production, as potential targets for reducing HIV associated chronic inflammation. sCD40-L is also produced by activated T-cells. Our results show an association between increased T-cell

activation (sCD25/sCD27 and PC1) and HCLD, which is consistent with previous reports implicating peripheral T-cell activation with HIV-associated pulmonary dysfunction (20). Soluble CD25 (IL2-RA) correlates well with surface CD25 expression (31) and is a marker of activated T-regulatory cells (CD25+) (32). Soluble CD25 has been used as a marker of disease severity in a number of inflammatory conditions (33) and is associated with an expanded Th17 response (34), which recruits pathogenic T cells to sites of inflammation in inflammatory diseases such as asthma (35). In non-atopic asthma patients sCD25 is negatively associated with FEV1. These patients typically experience broncho-obstructive reactions to inflammatory stimuli similar to HCLD which normalises with sCD25 (36). Intriguingly several of these markers (sCD40-L, CD25, CD27) were not associated with FEV1 z-score in the HCLD group. These findings could suggest that different pathways may be involved in the initiation and progression of HCLD, which warrants further study. Overall, these results suggest that treatment strategies which reduce the levels of platelet and T-cell activation may be beneficial for the prevention of HCLD.

We report an association between HCLD and 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 FEV1 decline at the population level (37). 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 (38). In chronic obstructive pulmonary disease (COPD), tissue injury is thought to be promoted by IFN-γ through release of MMP from activated macrophages (39). 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 HCLD. 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 (40), and is particularly marked in perinatal infection (41). Soluble 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 (42). 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 (43,44). D-Dimer is also strongly correlated with endothelial dysfunction, microbial translocation and sCD14 (45–47) and strongly supports evidence implicating platelets as drivers of pathology.

Owing to the essential role of the pulmonary extracellular matrix (ECM) for normal lung function, the association of several markers of extracellular matrix degradation (MMP1, 7 & 10) with HCLD is of interest (48). The level of MMPs from bronchoalveolar lavage (BAL) fluid samples have been associated with radiological markers of small airway disease and emphysema severity (49) with MMP-7 shown to promote pulmonary fibrosis (50). MMP-10 is expressed by multiple cell types in response to infection (51), and likely represents

increased immune activation in the HCLD group. MMP-1 is elevated in subjects with COPD and children with pulmonary TB (49) but is responsive to TB treatment (52).

As soluble biomarkers do not act in isolation, we sought to study the relationships between biomarkers in cases and controls to understand better the activated pathways that may drive pathology. Comparing how networks change between disease states can help understand processes involved in pathology We suggest that biomarkers of high centrality in the HCLD group offer the best potential for therapies aimed at reducing systemic immune activation. The increased centrality of the well described immune activation markers β2M, IFN-γ and IP-10 in cases is indicative of HCLD being associated with an elevated inflammatory environment. Interestingly, the centrality of ANG-1 was inverted, concordant with positive effects of high ANG-1 in persons living with HIV (53). However, this is not consistent with increased odds of HCLD in individuals with higher levels of ANG-1. Further work is required to understand the association between angiogenesis and HCLD. Interestingly the HCLD- network contains more negative correlations between biomarkers, likely representative of lower overall levels of immune activation and maintenance of functional regulation which becomes dysregulated in the context of HCLD.

There are several limitations to this study. The cross-sectional design means that the direction of temporal relationships is unknown. Ward *et al* found no relationship between bronchoalveolar lavage (BAL) and blood sCD14 (33), indicating that the levels of plasma soluble biomarkers may not represent local levels in relevant organs. High-resolution computed tomographic scans and BAL sampling from BREATHE participants would allow us better to describe the phenotype of HCLD and describe local inflammation within the

cohort. Due to the exploratory nature of the study, we have not corrected for multiple testing and results should be interpreted accordingly.

In conclusion, this study furthers the understanding of pathways associated with HCLD in children and adolescents living with PHIV and suggests that studies aimed at characterising T-cell activation alongside understanding the activity of platelets may be warranted. Furthermore, these results show that soluble biomarkers representing a wide array of pathways are associated with HCLD, and highlight a central role for sCD40-L, which itself showed good predictive ability for HCLD. These results act as a first step towards understanding the pathology of HCLD, alongside highlighting potential targets for therapeutic modalities that may have utility in prevention.

Acknowledgements: This work was supported by Research Council of Norway, through their GLOBVAC programme, with additional support from Helse Nord. RAF is funded by the Wellcome Trust through a Senior Fellowship in Clinical Science. DHB is funded by the Wellcome Trust PhD Programme in Genomic Medicine and Statistics, Grant number 108861/B/15/Z. AMR & VS are additionally funded by the UK Medical Research Council (MRC) and the UK department for international Development (DFID) under the MRC/DFID Concordant agreement, also part of the EDCTP2 programme supported by the European Union, Grant Ref: MR/R010161/1.

This research was funded in whole, or in part, by the Wellcome Trust.

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

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

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

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

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

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

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

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

14. Gonzalez-Martinez C, Kranzer K, McHugh G, Corbett EL, Mujuru H, Nicol MP, et al. 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. Hameiri Bowen D, Sovershaeva E, Charlton B, Schive C, Odland J, McHugh G, et al. Cytomegalovirus-Specific Immunoglobulin G is associated with chronic lung disease in children and adolescents from sub-Saharan Africa with perinatal HIV infection. Clin Infect Dis [Internet]. 2020 Nov 26; Available from: https://doi.org/10.1093/cid/ciaa1757

25. André P, Nannizzi-Alaimo L, Prasad SK, Phillips DR. Platelet-derived CD40L: The switch-hitting player of cardiovascular disease. Circulation. 2002;106(8):896–9.

26. Tabuchi A, Kuebler WM. Endothelium-platelet interactions in inflammatory lung disease. Vascul Pharmacol [Internet]. 2008;49(4–6):141–50. Available from: http://dx.doi.org/10.1016/j.vph.2008.06.004

27. 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.

28. Tennert K, Schneider L, Bischof G, Korn K, Harrer E, Harrer T, et al. Elevated CD40 ligand silences α interferon production in an HIV-related immune reconstitution inflammatory syndrome. AIDS [Internet]. 2013;27(2). Available from: https://journals.lww.com/aidsonline/Fulltext/2013/01140/Elevated_CD40_ligand_silences___interferon.21.aspx

29. Sui Z, Sniderhan LF, Schifitto G, Phipps RP, Gelbard HA, Dewhurst S, et al. Functional Synergy between CD40 Ligand and HIV-1 Tat Contributes to Inflammation: Implications in HIV Type 1 Dementia. J Immunol. 2007;178(5):3226–36.

30. Davidson DC, Jackson JW, Maggirwar SB. Targeting platelet-derived soluble CD40 ligand: A new treatment strategy for HIV-associated neuroinflammation? J Neuroinflammation. 2013;10:1–11.

31. 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.

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

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

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

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

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

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

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

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

40. 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.

41. 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.

42. 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).

43. 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.

44. 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

45. 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.

46. 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.

47. 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.

48. 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

49. 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.

50. 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.

51. 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.

52. 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.

53. Gulhati V, Soo J, Ransy DG, Brophy J, Kakkar F, Bitnun A, et al. Brief Report: Higher Levels of Angiopoietin-1 Are Associated with Early and Sustained Viral Suppression in Children Living with Vertically Acquired HIV. J Acquir Immune Defic Syndr. 2019;80(5):590–5.

Table 1: Clinical, demographic and anthropometric characteristics of participants included 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)	.833
Female, N (%)	46 (62.2)	166 (49.4)	.063
Zimbabwe, N (%)	55 (74.3)	241 (71.7)	.758
Malawi, N (%)	19 (25.7)	95 (28.3)s	-
Height-for-Age z-score, Median (IQR)	-1.5 (1.0)	-2.1 (1.2)	< .001
Stunted, N (%)	22 (29.7)	168 (50.0)	.002
Weight-for-Age z-score, Median (IQR)	-1.1 (1.2)	-2.2 (1.5)	< .001
Wasting, N (%)	14 (18.9)	176 (52.4)	< .001
$CD4^+$ T Cell Count, >350 Cells /mm^3, N	57 (78.1)	252 (75.2)	.715
HIV Viral Control (<200 copies/ml), N	38 (51.4)	145 (43.2)	.248
Log-10 HIV Viral Load Copies/ml,	2.0 (1.8)	2.6 (2.5)	.087
FEV1 Z-Score, Median (IQR)	0.6 (0.5)	-2.0 (0.7)	< .001
FEV1/FVC z-score, Mean (SD)	0.2 (0.8)	-0.7 (1.1)	< .001
FEV % Predicted, Mean (SD)	107.7 (6.1)	72.8 (9.9)	< .001
Duration ART in Years, Median (IQR)	6.5 (2.8)	6.4 (3.2)	.758
Ever Treated for TB, N (%)*	9 (12.2)	97 (29.9)	.005
First Line Regimen - ATV/LPV/PI, N (%)	66 (89.2)	249 (74.1)	.009
Second Line Regimen - EFV/NVP, N (%)*	8 (10.8)	87 (25.9)	-

IQR = Interquartile Range, N= Number, ATV = Atazanavir, LPV=Lopinavir, PI= Protease inhibitor EFV= Efavirenz NVP = Nevirapine, *One data point missing, ** Two data points missing

Table 2. Biomarkers and Principal Components associated with HCLD

	Univariate OR (95% CI, P)	Adjusted OR (95% 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)
PC1	1.60 (1.39-1.85, p< .001)	1.54 (1.33-1.80, p<0.001)
PC3	1.39 (1.15-1.70, p= .001)	1.44 (1.16-1.81, p=0.001)
PC4	0.68 (0.56-0.83, p< .001)	0.62 (0.49-0.77, p<0.001)
PC5	0.77 (0.62-0.96, p= .022)	0.71 (0.54-0.92, p=0.012)
PC6	0.77 (0.61-0.96, p= .020)	0.74 (0.58-0.94, p=0.014)
PC7	1.28 (1.00-1.64, p= .053)	1.47 (1.11-1.95, p=0.007)

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 3: Biomarkers associated with lung function in HCLD+ participants

	Univariate β ± SE	P-Value	Adjusted β ± SE	P-Value
MMP-8	-0.122 ± 0.039	.002	-0.097 ± 0.039	.012
MMP-10	-0.161 ± 0.038	<.001	-0.132 ± 0.04	.001
ANG-1	0.08 ± 0.039	.041	0.077 ± 0.039	.046
CRP	-0.149 ± 0.038	<.001	-0.128 ± 0.038	.001
IP-10	-0.089 ± 0.039	.023	-0.08 ± 0.04	.048
E-Selectin	-0.104 ± 0.039	.008	-0.082 ± 0.039	.037
Fas	0.083 ± 0.039	.035	0.082 ± 0.039	.033
GCSF	-0.106 ± 0.039	.007	-0.11 ± 0.039	.006
VCAM-1	-0.105 ± 0.039	.007	-0.09 ± 0.039	.021
VEGF	-0.117 ± 0.039	.003	-0.113 ± 0.039	.004
PC1	-0.053 ± 0.017	.002	-0.044 ± 0.018	.014
PC4	0.084 ± 0.03	.005	0.096 ± 0.03	.001
PC5	-0.105 ± 0.033	.002	-0.092 ± 0.036	.011
PC7	0.083 ± 0.038	.029	0.082 ± 0.038	.031

Linear Regression of $\log_{10}$ scaled biomarkers and FEV1 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. β = coefficient

Figure 1: Comparison of biomarkers in cases and controls

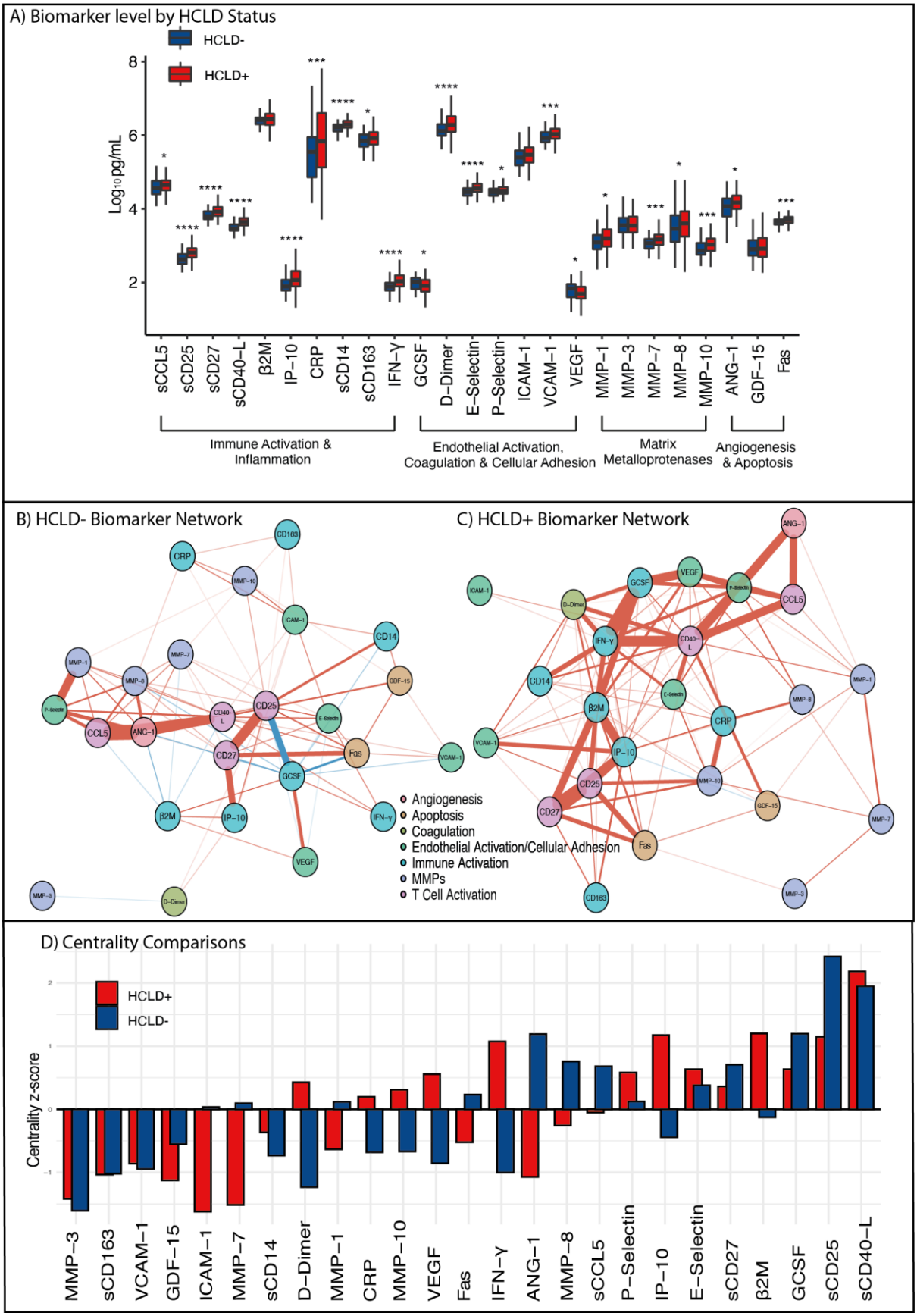
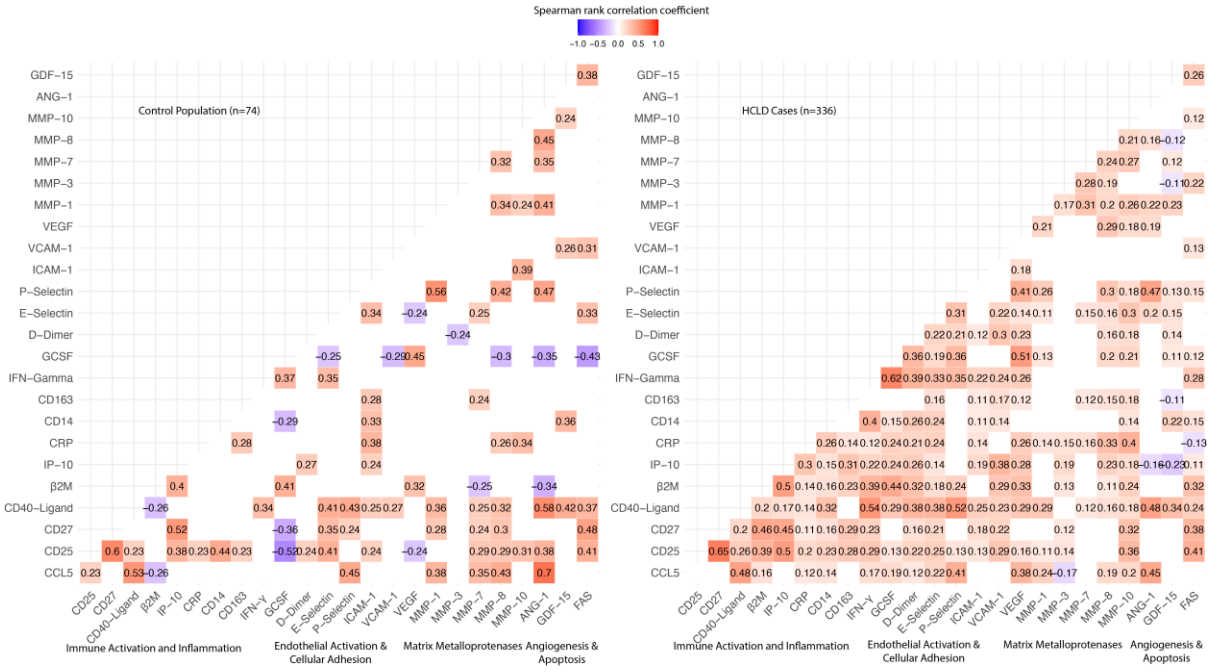

A) Biomarker level by HCLD Status
HCLD-
HCLD+
Log10 pg/mL
sCCL5
sCD25
sCD27
sCD40–L
β2M
IP–10
CRP
sCD14
sCD163
IFN–γ
GCSF
D–Dimer
E–Selectin
P–Selectin
ICAM–1
VCAM–1
VEGF
MMP–1
MMP–3
MMP–7
MMP–8
MMP–10
ANG–1
GDF–15
Fas
Immune Activation & Inflammation
Endothelial Activation, Coagulation & Cellular Adhesion
Matrix Metalloprotenases
Angiogenesis & Apoptosis
B) HCLD- Biomarker Network
C) HCLD+ Biomarker Network
Angiogenesis
Apoptosis
Coagulation
Endothelial Activation/Cellular Adhesion
Immune Activation
MMPs
T Cell Activation
D) Centrality Comparisons
HCLD+
HCLD-
Centrality z-score

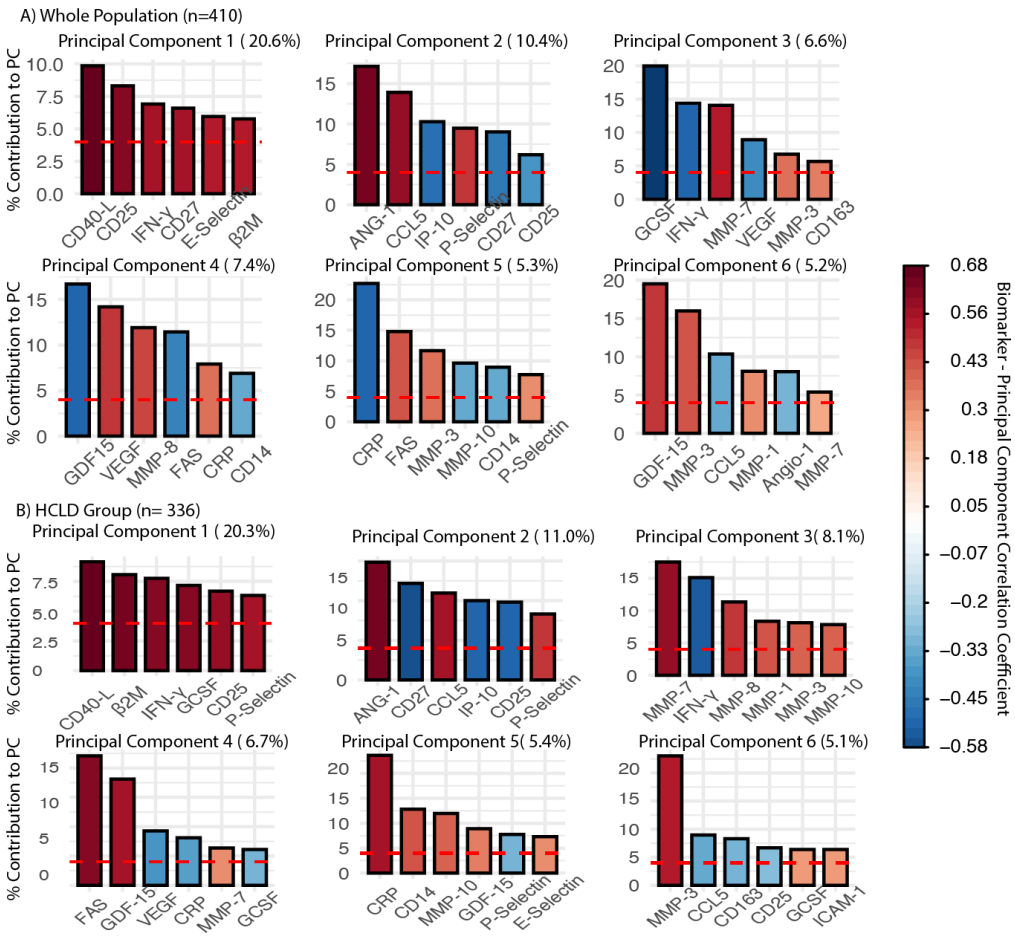
Figure 1 A) Comparison of Log_{10} biomarker level between cases (HCLD+) and controls (HCLD-). Levels compared by Wilcoxon rank sum test. Stars represent significance level at ≤ .05, ≤ .01, ≤ .001 and ≤ .0.0001 respectively. B/C) Network plots showing strength and direction of correlations between biomarkers in HCLD+ and HCLD- individuals. Colour saturation and width of edges correspond to absolute weight and scale relative to strongest weight in the graph. Nodes arranged by spring format. Biomarkers coloured by *a priori* biological pathway. Only significant ($p < .05$) correlations $r^2 > 0.2$ are shown D) Node strength centrality z-scores for nodes in the networks. Score indicates the relative interconnectedness of each biomarker within the network.

Supplementary Figure 1: Biomarker Correlations



Spearman rank correlations between biomarkers in A) HCLD- controls B) HCLD+ cases. Number within square indicates spearman rank correlation coefficient between biomarkers. Blank squares represent no significant correlation between biomarker pairs ($p > .05$).

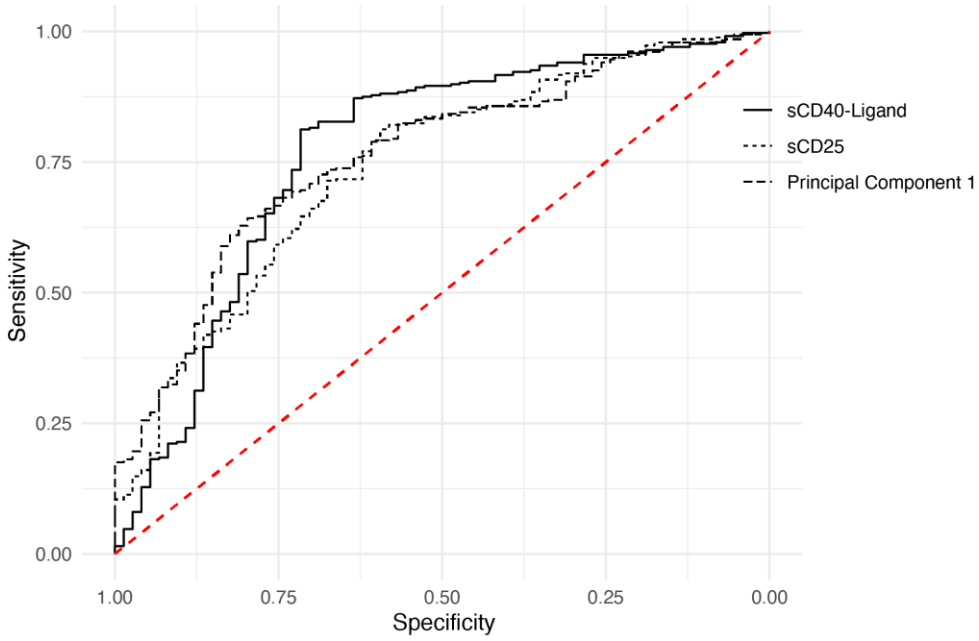
Supplementary Figure 2: Contribution of biomarkers principal components



Contribution of individual biomarkers to principal components derived from the A) Whole population and B) Participants with HCLD. Percentages in brackets show how much of

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

Supplementary Figure 3: Sensitivity and Specificity of markers for HCLD



ROC curve for top three performing variables. Only variables with AUC >0.7 included

Diagonal line represents no predictive ability. AUC = Area under the Curve