ERJ open research
THE BEST IN OPEN ACCESS BASIC, TRANSLATIONAL & CLINICAL RESPIRATORY RESEARCH

Early View

Original article

High serum G-CSF characterises neutrophilic COPD exacerbations associated with dysbiosis

Arindam Chakrabarti, Jordan S. Mar, David F. Choy, Yi Cao, Nisha Rathore, Xiaoying Yang, Gaik W. Tew, Olga Li, Prescott G. Woodruff, Christopher E. Brightling, Michele Grimbaldeston, Stephanie A. Christenson, Mona Bafadhel, Carrie M. Rosenberger

Please cite this article as: Chakrabarti A, Mar JS, Choy DF, *et al*. High serum G-CSF characterises neutrophilic COPD exacerbations associated with dysbiosis. *ERJ Open Res* 2021; in press (https://doi.org/10.1183/23120541.00836-2020).

High serum G-CSF characterizes neutrophilic COPD exacerbations associated with dysbiosis

Arindam Chakrabarti [1], Jordan S. Mar [1], David F. Choy [1], Yi Cao [2], Nisha Rathore [1], Xiaoying Yang [3], Gaik W. Tew[4], Olga Li [4], Prescott G. Woodruff [5], Christopher E. Brightling [6], Michele Grimbaldeston [4], Stephanie A. Christenson [5], Mona Bafadhel [7], Carrie M. Rosenberger [1]*

Affiliations:

[1.] Biomarker Discovery OMNI, [2.] Bioinformatics, [3.] Biostatistics, [4.] OMNI Biomarker Development, Genentech, Inc. 1 DNA Way, South San Francisco, CA 94080; [5.] Department of Medicine, UCSF, San Francisco, California, USA. [6.] Department of Respiratory Sciences, University of Leicester, Leicester, UK;

[7.] Nuffield Department of Medicine, University of Oxford, UK;

*** Correspondence:** Carrie M. Rosenberger, Biomarker Discovery OMNI, Genentech, Inc. 1 DNA Way, South San Francisco, CA 94080 E-mail: rosenberger.carrie@gene.com

Summary (230 characters)

Non-invasive biomarkers to characterize COPD exacerbation subtypes are limited. High serum G-CSF enriches for COPD exacerbations associated with bacterial infection and neutrophilic inflammation preceded by lung microbial dysbiosis.

Word count: Abstract: 247 Manuscript: 3000

ABSTRACT

Introduction: COPD exacerbations are heterogeneous and can be triggered by bacterial, viral, or non-infectious insults. Exacerbations are also heterogeneous in neutrophilic or eosinophilic inflammatory responses. A non-invasive peripheral biomarker of COPD exacerbations characterized by bacterial/neutrophilic inflammation is lacking. G-CSF is a key cytokine elevated during bacterial infection and mediates survival, proliferation, differentiation, and function of neutrophils.

Objective: We hypothesized that high peripheral G-CSF would be indicative of COPD exacerbations with a neutrophilic and bacterial phenotype associated with microbial dysbiosis.

Methods: Serum G-CSF was measured during hospitalized exacerbation and after 30 days of recovery in 37 subjects. In a second cohort, serum and sputum cytokines were measured in 59 COPD patients during stable disease, at exacerbation, and 2- and 6-weeks following exacerbations.

Results: Serum G-CSF is increased during exacerbation in a subset of patients. These exacerbations were enriched for bacterial and not viral or type-2 biologies. The median serum G-CSF levels was 1.6-fold higher in bacterial exacerbation compared to non-bacterial exacerbation (22 pg/ml vs 13 pg/ml, p = 0.0007). Serum G-CSF classified bacterial exacerbations with an AUROC=0.76. Exacerbations with a ≥2-fold increase in serum G-CSF were characterized by neutrophilic inflammation, with increased sputum and blood neutrophils and high sputum IL-1β, IL-6, and SAA. These exacerbations were preceded by dysbiosis, with decreased microbiome diversity and enrichment of respiratory pathogens such as *Haemophilus* and *Moraxella*. Furthermore, serum G-CSF at exacerbation classified neutrophilic-dysbiotic exacerbations (AUROC=0.75).

Conclusions: High serum G-CSF enriches for COPD exacerbations characterized by neutrophilic inflammation with underlying bacterial dysbiosis.

INTRODUCTION

Acute exacerbations of chronic obstructive pulmonary disease (COPD) decreases patient quality of life, contributes to progressive airway function decline, and can be fatal [1], [2]. Exacerbations are defined by an acute worsening of respiratory function leading to therapeutic intervention with antibiotics or corticosteroids [2]. There have been multiple recent efforts to define exacerbations more precisely by their molecular features, which has revealed heterogeneity in this key clinical event [3–5]. The identification of biomarkers to classify exacerbations is a step towards developing a mechanistic understanding of the underlying inflammatory etiology, which is needed for the development of targeted therapies.

Unsupervised clustering analysis has previously identified four exacerbation phenotypes: bacterial, viral, eosinophilic, and pauci-inflammatory. These clusters were defined by pathogens, eosinophils or little inflammation respectively, in sputum collected during exacerbation [3]. The most sensitive and specific inflammatory biomarkers of these exacerbation phenotypes were sputum IL-1β (bacterial-predominant), serum CXCL10 (IP-10, viral-predominant), and the percentage of blood eosinophils (eosinophil-predominant). A similar clustering approach in COPD and asthma patients that did not include microbiological information yielded a neutrophil-predominant cluster containing features attributed to both bacterial and viral exacerbations: high sputum % neutrophils, IL-1β, TNFα, IL-8 (bacterial), and serum IP-10 (viral) [4]. The subset of exacerbations characterized by high sputum IL-1β and presence of pathogenic bacteria such as *Haemophilus influenzae* are also characterized by increased sputum levels of the neutrophil-promoting cytokines G-CSF and IL-6 [6]. Neutrophils are the predominant inflammatory cell in respiratory samples during bacterial and viral exacerbations in many patients, and a potential therapeutic target [7, 8]. Sputum neutrophil elastase activity is a biomarker of bacterial exacerbation [9], an etiology accounting for approximately half of COPD exacerbations [3]. A recent report further sub-grouped neutrophilic COPD based on the abundance of *Haemophilus* [10]. Since sputum cannot be produced by all patients and is not routinely collected during exacerbations, the present study focuses on identifying a peripheral marker that can be indicative of a COPD exacerbation phenotype associated with underlying airway dysbiosis with overabundance of potentially pathogenic airway bacteria and high sputum neutrophils.

We hypothesized that peripheral G-CSF could be a biomarker of a neutrophilic COPD phenotype and would enrich for bacterial/neutrophilic exacerbations. Serum G-CSF correlates with airway G-CSF and the best non-invasive biomarker of lung neutrophils in the bronchoalveolar lavage fluid of patients with acute respiratory distress syndrome (ARDS) [11–13]. Here, we show that increased serum G-CSF classifies COPD exacerbations associated with airway dysbiosis enriched for potentially pathogenic bacteria and a neutrophil-predominant inflammatory response.

METHODS

Study Populations

Clinical studies were approved by ethics review committees. The ethics number for the BEAT:COPD study is 07/H0406/157.

LEUKO cohort. Biomarkers were measured in 54 samples at exacerbation (Day 0) and 52 samples at recovery (Day 30) with 37 paired measurements from a double-blind, placebo-controlled, parallel group study of COPD subjects randomized to receive either Zileuton tablets 600 mg or placebo tablets orally, 4 times per day for 14 days (study described in [14], NCT00493974). Inclusion criteria included an admitting diagnosis of AECOPD (defined as an acute increase in dyspnea, sputum volume, and/or sputum purulence without other attributable cause), ≥10 pack-years smoking history, and an FEV1 < 60% predicted at time of inclusion or an inability to perform spirometry due to dyspnea.

MRC cohort. Biomarkers were measured in 74 exacerbation samples from 59 subjects with 56 paired serum and sputum measurements (Supplementary Table 1) along with preceding stable and post-exacerbation samples from a prospective observational study [3]. Inclusion criteria included physician diagnosed COPD, post-bronchodilator FEV1/FVC ratio < 0.7, and ≥ 1 exacerbation in the preceding 12 months. Samples were collected every 3 months during stable disease for 1 year. Baseline and stable visits were 8-weeks free of an exacerbation that required treatment with oral corticosteroid or antibiotics. Samples were collected during exacerbation

and at 2 and 6-week follow up visits. Subjects received standard of care oral steroid and/or antibiotic treatments for exacerbations [3].

Classification of exacerbation subtype

Bacteria-, virus-, and sputum eosinophil–associated COPD exacerbations are defined according to the original publication from MRC cohort [3]. Briefly, bacterial exacerbations were defined as a positive bacterial pathogen on routine culture (*Haemophilus influenzae, Moraxella catarrhalis, Streptococcus pneumoniae, Staphylococcus aureus* or *Pseudomonas aeruginosa*) and/or a total aerobic colony forming unit count (CFU) $\geq 10^7$. Viral exacerbations were defined as one that had a positive sputum viral PCR. Eosinophilic exacerbations were defined as the presence of >3% non-squamous cells following sputum cytospin slide preparation. Neutrophilic COPD exacerbations were classified based on sputum neutrophil $\geq$ 61% at exacerbation, and subdivided by the combined relative abundance of *Haemophilus* and *Moraxella* (dysbiotic >0.41 and balanced $\leq$0.41*)* as defined in [10].

Measurements of biomarkers and statistical methods

Serum and sputum biomarkers were measured using multiplex ELISA platform including Luminex and as previously described [3]. Statistical analyses were performed using PRISM version 8 (Graph-PAD PRISM, La Jolla, CA) or R-programming software. Biomarker levels were log transformed and medians with 95% confidence intervals (CI) shown . Statistical significance between exacerbation and recovery or preceding stable visits was calculated by Wilcoxon matched-pairs signed rank test. Statistical significance between G-CSF high vs low groups or exacerbations +/- bacterial infection was calculated using Mann-Whitney U test. Area under the receiver operating characteristic curves were calculated for bacteria- versus non-bacteria-associated exacerbation. Correlation heatmaps and the pairwise Spearman correlation coefficients were computed using the “cor” command in R, and plots generated using “ggplot2” and “ggpubr” R packages.

16SV3-V5 rRNA Sequence Data Processing

Raw sequence data was obtained from the NCBI Sequence Read Archive database under accession number SRP065072 and processed using QIIMEv1.9 [15]. Raw sequence data was first demultiplexed and quality filtered, keeping reads between 200-1000bp with a minimum average q-score >29 and zero ambiguous base calls. High quality reads were then binned into Operational Taxonomic Units (OTUs) at >97% sequence identity using *sortmerna* and the Greengenes bacterial rRNA database (v13_8) for closed OTU picking followed by open OTU picking using *sumaclust* [16]. Chimeric OTUs were removed with ChimeraSlayer (http://microbiomeutil.sourceforge.net/). Chimera-free OTU abundance tables were then rarefied to 4233 reads per sample to account for variable sequencing depths.

Microbiome Analysis

α- and β-diversity measures were calculated in QIIMEv1.9. All statistical analysis was conducted in the R statistical environment (https://cran.r-project.org/). A Wilcoxon rank-sum test was performed to compare α-diversity between exacerbation types. Bray-Curtis distance matrices were visualized via NMDS and permutational multivariate analysis of variance (PERMANOVA) was used to determine statistical relationships between metadata (i.e. experimental group) and bacterial microbiota composition in the R statistical environment using the vegan package (https://github.com/vegandevs/vegan). To identify significantly enriched or depleted bacterial taxa in relevant experimental groups, the DESeq2 R-package was used as described by McMurdie et al [17]. Resulting p-values were adjusted for false discoveries using the Benjamini-Hochberg correction [18].

Human bronchial epithelial cell culture and treatment

Normal bronchial epithelial cells (Lonza, Basel, Switzerland) were grown and differentiated in air-liquid interface prior to treatment with recombinant human IL-1β (10 ng/ml) (R&D Systems Inc. Minneapolis, MN; Cat # 201-LB-010) for 24 hours. Cytokines were measured in basal supernatants by multiplex Luminex (EMD Milliplex Cat. No # HCYTMAG-60K-PX30). Total RNA was isolated using RNeasy Mini Kit (Qiagen, Germantown, MD) and sent for standard RNA sequencing (Illumina, single end read, 50 bp, 30M).

RESULTS

Serum G-CSF is significantly elevated in a subset of hospitalized exacerbations

We measured serum G-CSF during severe hospitalized exacerbation and after 30 days of recovery in available samples from the LEUKO study (cohort demographics are summarized in Table 1) [14]. Serum G-CSF was increased at the time of hospitalization compared with the 30-day recovery sample ($p<0.0001$, figure 1A). A strong correlation was observed between the absolute level of serum G-CSF at exacerbation and the fold change between exacerbation and recovery samples (Spearman correlation=0.8, $p<0.0001$, supplementary figure 1A).

A cutoff of 2-fold difference between G-CSF levels at exacerbation and recovery identified 21/37 G-CSF^{high} exacerbations with a 2.5-fold higher median serum G-CSF level (figure 1B, $p<0.0001$). The fold change of G-CSF was correlated with fold change of other biomarkers of neutrophilic inflammation such as IL-8 and SAA1 (figure 1C). In addition, we observed that the G-CSF^{high} group tended to respond better to antibiotics treatment with a shorter length of hospitalization compared to G-CSF^{low} groups of patients (supplementary figure 1 B-C), although it did not achieve statistical significance.

Elevated serum G-CSF levels identifies exacerbations characterized by pathogenic bacteria and neutrophilic inflammation

To better characterize G-CSF^{high} exacerbations, we utilized a second available COPD exacerbation cohort which offered paired serum and sputum samples, richer longitudinal sampling and available blood and sputum cell counts. This cohort of mild to moderate exacerbations pioneered the classification of exacerbation phenotypes using unbiased clustering analysis [3]. We measured G-CSF in 74 primary exacerbation samples that were preceded by a sample collected at a stable disease visit within the preceding 3 months, and at 2 weeks and 6 weeks post-exacerbation (see table 1 for patient demographics). G-CSF levels were unaffected by smoking status or FEV_1 ($p>0.05$ Mann-Whitney). Increased serum G-CSF was observed in a subset of exacerbations ($p<0.0001$, Figure 2A). To characterize these G-CSF ^{high} exacerbations, we classified exacerbations by previously published criteria of eosinophilic, bacterial or viral [3] (supplementary table 1) and observed significantly increased serum G-CSF in non-eosinophilic and bacterial exacerbations and not in eosinophilic nor viral exacerbations

(Figures 2B-E). As bacterial and viral co-infections occur during COPD exacerbations (supplementary table 1) [3], we also examined the subset of exacerbations with any diagnosed pathogenic bacteria or virus and also observed increased G-CSF ($p<0.0001$, Figure 2F).

We then characterized the inflammatory characteristics of the twenty G-CSF^{high} exacerbations, defined by ≥2-fold change in serum G-CSF between exacerbation and preceding stable visit. G-CSF^{high} exacerbations showed a greater increase in median blood neutrophil counts than G-CSF^{low} exacerbations (4.8-fold vs 1.5-fold increase) (Figure 3A). These G-CSF^{high} exacerbations showed a 7.5-fold higher median level of blood neutrophils and higher levels of IL-6 and serum amyloid A1 (SAA1), which promote neutrophilic inflammation (Figure 3B-D). Serum levels of other cytokines associated with bacterial (IL-8, TNF-α, IL-1β), viral (CXCL10), and type 2 inflammation (blood eosinophil counts) did not differ between G-CSF^{high} and G-CSF^{low} exacerbations ($p>0.05$, data not shown). G-CSF^{high} exacerbations were also characterized by an increased sputum neutrophil percentage and sputum IL-6 (Figure 3E-F). Serum G-CSF at exacerbation was well correlated with multiple metrics of neutrophilic inflammation: blood and sputum neutrophils, serum IL-6, SAA1, CRP (Spearman correlations rho≥0.5), and moderately correlated with sputum IL-6 and IL-8 (Spearman correlations rho≥0.3), and not biomarkers of viral or eosinophilic inflammation (Figure 3G and supplementary figure S2). When fold change between stable disease and exacerbation were compared, serum G-CSF correlated with blood neutrophils, serum IL-6 and SAA1, and sputum IL-6 (Spearman correlations rho≥0.4, Figure 3H and supplementary Figure S2). Calculating fold change between levels at exacerbation and either baseline or 6-week recovery samples, in line with the results in the LEUKO cohort, gave comparable results (data not shown).

The identification of biomarkers that could be measured during stable disease to predict exacerbation subtypes would enable patient enrichment, but have been challenging to identify. In a previous study of this cohort, no single biomarker measured in stable disease could predict an exacerbation phenotype well. We observed a trend that patients with higher G-CSF during stable disease were more likely to have exacerbations characterized by neutrophilic inflammation. When subjects with the highest and lowest quartile of stable disease G-CSF were compared, high stable disease G-CSF was associated with higher blood

neutrophils and significantly higher sputum neutrophils, sputum IL-8 and serum G-CSF at exacerbation ($p<0.05$) (supplementary figure 3).

Elevated serum G-CSF correlates with sputum IL-1β, a key biomarker of bacterial exacerbations

Bacterial exacerbations exhibited a larger increase in G-CSF and higher absolute levels when compared with non-bacterial exacerbations (Figure 4A-B). Sputum IL-1β is the strongest immune correlate of bacterial exacerbations [3][4, 6] (Figure 4C). Moreover, a strong correlation of sputum IL-1β and serum G-CSF levels at exacerbation was observed (Spearman rho=0.6, $p<0.0001$, Figure 4D). A previous study using this cohort identified CRP as the best serum predictor of bacterial exacerbation [3]. In the subset of samples included in this study, we observed the expected higher serum levels of CRP in bacterial exacerbations and strong correlation between CRP and G-CSF (Spearman correlation r=0.7, $p<0.0001$, Figure 4E-F). Performance of G-CSF in classifying bacterial exacerbations was evaluated by calculating the area under the receiver operating characteristics curve, and compared with blood neutrophil count and serum CRP. Serum G-CSF classified bacterial exacerbations with an AUROC=0.76 which was comparable to AUROC of blood neutrophil and serum CRP (0.70 and 0.71 respectively) (Figure 4G).

G-CSF^{high} exacerbations are preceded by altered sputum microbiome

COPD exacerbations are associated with lung microbiome dysbiosis [19–25]. Exacerbation subtype cluster characterized by high sputum IL-1β and mediators of neutrophilic inflammation are enriched by the presence of potentially pathogenic bacteria such as *Haemophilus* and *Moraxella* [6, 21, 24, 25]. We hypothesized that bacterial exacerbations and exacerbations characterized by high serum G-CSF levels are preceded by an altered lung microbiome composition. To test this hypothesis, microbiome analysis was performed on available sputum from exacerbations (n=41) and the preceding stable visits (n=46).

Reduced α-diversity and a gross shift in overall sputum microbiome composition was observed prior to bacterial exacerbations (Figure 5A-B). Moreover, relative abundance of potential respiratory pathogens *Haemophilus* and *Moraxella* displayed greater than 7.0- and

1.5-fold enrichment, respectively, prior to bacterial exacerbations compared to non-bacterial exacerbations, though the increase in *Moraxella* did not meet statistical significance after correction for FDR (Figure 5C-D). In general, these observations persisted through exacerbation visits with reduced α-diversity and *Haemophilus* and *Moraxella* enrichment also occurring in sputum samples collected during bacterial exacerbations (supplementary Figure 4A-D). Although α-diversity was not significantly reduced during stable visits preceding G-CSF^{high} exacerbations (Figure 5E), an altered sputum microbiome composition, characterized by greater than 13.0-fold enrichment of both *Haemophilus* and *Moraxella (*with only the increase in *Moraxella* remaining significant after FDR correction), was still observed prior to these G-CSF^{high} exacerbations (Figure 5F-H). However, while *Moraxella* enrichment persisted in sputum collected during G-CSF^{high} exacerbations, *Haemophilus* abundance was slightly reduced (supplementary Figure 4E-H).

High serum G-CSF levels at exacerbation characterizes a neutrophilic-dysbiotic phenotype

A recent report has shown that two major types of airway ecology exists for neutrophilic COPD, differentiated by the predominance of *Haemophilus* [10]. Accordingly, we examined whether high serum G-CSF levels were associated with the exacerbation endotype characterized by high sputum neutrophil and overabundance of potentially pathogenic organisms such as *Haemophilus* and/or *Moraxella*. We found that serum G-CSF and blood neutrophil counts but not serum CRP were significantly higher in Neutrophilic COPD exacerbations, defined by sputum neutrophil $\geq$61% (Figure 6A-C). Next, we stratified the Neutrophilic COPD exacerbations into Neutrophilic-dysbiotic and Neutrophilic-balanced exacerbations based on combined relative abundance of *Haemophilus* and *Moraxella* $\geq$ 0.41. Median serum G-CSF was significantly higher in Neutrophilic-dysbiotic COPD exacerbations compared to the Neutrophil^{Low} COPD exacebations, while blood neutrophil counts and serum CRP levels were not. Median G-CSF level trended higher in Neutrophilic-dysbiotic exacerbations compared to Neutrophilic-balanced exacerbations but was not significant (p=0.07) (Figure 6D-F). Performance of G-CSF in classifying Neutrophilic-dysbiotic exacerbations was evaluated by calculating the area under the receiver operating characteristics curve and compared with blood neutrophil and serum CRP. Serum G-CSF stratified Neutrophilic-dysbiotic COPD with an

AUROC=0.75 while blood neutrophil counts and serum CRP had AUROC of 0.50 and 0.59 respectively (Figure 6G).

Treatment of primary lung epithelial cells with IL-1β induces G-CSF production

We were interested in the mechanistic connection between the module of neutrophil-associated biomarkers correlating with G-CSF. An IL-1β response signature in lung epithelial cells was previously identified, which included G-CSF, IL-6 and TNFα [6]. We confirmed that IL-1β was sufficient to induce G-CSF and IL-6 secretion from primary bronchial epithelial cells cultured at air-liquid interface (Figure 7A-B). IL-1β also increased transcription of multiple neutrophil-recruiting chemokines (CXCL1, CXCL2, CXCL8/IL-8) and secretion of IL-8 (Figure 7D-G and 7C). However, transcription of IL-6 was not induced by IL-1β (Figure 7H). CRP is an acute phase protein released by the liver in response to IL-6, however, IL-1β did not induce transcription of CRP by lung epithelial cells (data not shown). Therefore, systemic G-CSF might be mechanistically connected to lung neutrophilic response to dysbiosis orchestrated by local production of IL-1β.

DISCUSSION

A better understanding of the molecular drivers of exacerbation phenotypes is needed to develop targeted therapeutic interventions for COPD patients. Biomarker classifiers are needed to study exacerbation heterogeneity in observational patient cohorts and to interpret data from clinical trials. This exploratory study identifies G-CSF as a novel molecular feature of the bacterial/neutrophilic exacerbation subtype described by other groups [3, 4, 6, 10]. Serum G-CSF classifies neutrophilic-dysbiotic exacerbations better than serum CRP or blood neutrophils. This recently described exacerbation subtype is characterized by high sputum neutrophils and IL-1β and abundance of potential bacterial pathogens such as *Haemophilus* and *Moraxella* by 16S sequencing (Supplementary Figure 5) [10]. Serum G-CSF correlates with neutrophils and neutrophil-promoting factors such as IL-6 and IL-8 in the lung, as sampled by induced sputum during exacerbation, and in the circulation. Serum G-CSF correlates with sputum IL-1β and we demonstrate using physiologically-relevant primary differentiated bronchial epithelial cells that IL-1β stimulation is sufficient for release of G-CSF, IL-6, and IL-8. G-

CSF has pleiotropic effects on neutrophils, which are a predominant inflammatory cell type in the lungs of many COPD patients [13, 26–28]. It promotes granulopoiesis and inhibits apoptosis, which together contribute to increased neutrophil numbers in the lung [26, 27, 29, 30]. G-CSF can also alter neutrophil functionality such as chemotaxis, ROS production and degranulation [31, 32]. How this may contribute to impaired response to infection and exacerbation severity in COPD patients is an interesting area for further investigation.

Predictors of exacerbation subtypes that could be measured during stable disease would greatly facilitate patient selection into clinical trials of targeted therapies. Our data suggests that the increased frequency of bacterial etiology in exacerbations associated with bacterial infection and G-CSF^{high} exacerbations may be related to a preceding dysbiosis in the lung microbiome. Although earlier reports highlighted microbial dysbiosis during exacerbation, this is the first instance where a preceding stable state is shown to be enriched with potentially pathogenic respiratory bacteria and associated with exacerbation subtypes defined by bacterial infection or high serum G-CSF. Moreover, subtypes of COPD are predicted to have distinct lung microbiome composition and stabilities over time [19], which highlights the importance of better understanding the relationship of the lung microbiome during stable disease to exacerbation phenotypes from both biomarker and mechanistic perspectives. We also observed a trend that the patients with high serum G-CSF measured during stable disease were more likely to have exacerbations characterized by neutrophilic inflammation (Supplementary Figure 3). The relationships between stable disease levels of serum G-CSF, lung neutrophilic inflammation and microbial dysbiosis and exacerbation phenotypes requires characterization in larger cohorts.

Limitations of this study include the small number of samples available for exacerbation subtypes and limited size of the microbiome dataset. Validation studies are underway to evaluate G-CSF in additional cohorts. The trend that patients with G-CSF^{high} exacerbations responded better to antibiotic therapy with shorter duration of hospitalization observed in the LEUKO cohort requires validation in larger independent cohorts. The MRC cohort did not sample all exacerbations, precluding an analysis of the stability of G-CSF^{high} exacerbations in patients and examination of prognostic biomarkers for the rate of G-CSF^{high} exacerbations. The

viral and bacterial PCR panels were not comprehensive and so viral and bacterial exacerbations are potentially under-diagnosed. Applying metagenomics approaches in future studies will yield a more comprehensive view of relationships between microbes and host inflammatory responses.

Cumulatively, our data suggest that COPD patients having bacterial exacerbation are associated with dysbiotic lung microbiome at a preceding stable state with lower microbial diversity driven by increased abundance of respiratory pathogens like *Haemophilus* and *Moraxella*. This exacerbation endotype is associated with an overabundance of bacterial load in the lung reflected by high levels of sputum IL-1β , neutrophils and serum G-CSF. Accordingly, confirming serum G-CSF as a prospective biomarker for neutrophilic-dysbiotic exacerbations in additional cohorts will aid in enriching subjects for precision medicine.

ACKNOWLEDGEMENTS

We thank John E. Connett, Prescott G. Woodruff and other CCRN collaborators for sharing the samples and data from LEUKO study. We thank Christopher E. Brightling and Mona Bafadhel for sharing samples and data from the MRC cohort. We thank Biorender for their help in the illustration.

FUNDING. This work was supported by Genentech, Inc.

CONFLICTS OF INTERESTS

AC, JSM, DFC, YC, NR, GWT, XY, OL, MG and CMR were employees of Genentech, Inc. during the execution of this study. PGW has received grant support from Genentech (outside the present work) and has been compensated for consulting activities by Sanofi, Regeneron, Theravance, Astra Zeneca, NGM Pharmaceuticals, Glenmark Pharmaceuticals, Amgen, also outside the present work. SAC, CEB and MB have no conflicts of interest.

References

1. Vogelmeier CF, Criner GJ, Martínez FJ, Anzueto A, Barnes PJ, Bourbeau J, Celli BR, Chen R, Decramer M, Fabbri LM, Frith P, Halpin DMG, Varela MVL, Nishimura M, Roche N, Rodríguez-Roisin R, Sin DD, Singh D, Stockley R, Vestbo J, Wedzicha JA, Agustí A. Erratum to "Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report: GOLD Executive Summary" (Arch Bronconeumol. (2017) 53 (128–49) (S0300289617300352) (10.1016/j.arbres.2017.02.001)). *Arch. Bronconeumol.* 2017; 53: 411–412.
2. Wedzicha JA, Wilkinson T. Impact of chronic obstructive pulmonary disease exacerbations on patients and payers. *Proc. Am. Thorac. Soc.* 2006; 3: 218–221.
3. Bafadhel M, McKenna S, Terry S, Mistry V, Reid C, Haldar P, McCormick M, Haldar K, Kebadze T, Duvoix A, Lindblad K, Patel H, Rugman P, Dodson P, Jenkins M, Saunders M, Newbold P, Green RH, Venge P, Lomas DA, Barer MR, Johnston SL, Pavord ID, Brightling CE. Acute exacerbations of chronic obstructive pulmonary disease: Identification of biologic clusters and their biomarkers. *Am. J. Respir. Crit. Care Med.* 2011; 184: 662–671.
4. Ghebre MA, Bafadhel M, Desai D, Cohen SE, Newbold P, Rapley L, Woods J, Rugman P, Pavord ID, Newby C, Burton PR, May RD, Brightling CE. Biological clustering supports both "dutch" and "british" hypotheses of asthma and chronic obstructive pulmonary disease. *J. Allergy Clin. Immunol.* 2015; 135: 63-72.e10.
5. Mathioudakis AG, Janssens W, Sivapalan P, Singanayagam A, Dransfield MT, Jensen J-US, Vestbo J. Acute exacerbations of chronic obstructive pulmonary disease: in search of diagnostic biomarkers and treatable traits. *Thorax* 2020; : thoraxjnl-2019-214484.
6. Damera G, Pham TH, Zhang J, Ward CK, Newbold P, Ranade K, Sethi S. A Sputum Proteomic Signature That Associates with Increased IL-1β Levels and Bacterial Exacerbations of COPD. *Lung* 2016; 194: 363–369.
7. Jurcevic S, Humfrey C, Uddin M, Warrington S, Larsson B, Keen C. The effect of a selective CXCR2 antagonist (AZD5069) on human blood neutrophil count and innate immune functions. *Br. J. Clin. Pharmacol.* 2015; 80: 1324–1336.
8. Lazaar AL, Miller BE, Tabberer M, Yonchuk J, Leidy N, Ambery C, Bloomer J, Watz H, Tal-

Singer R. Effect of the CXCR2 antagonist danirixin on symptoms and health status in COPD. *Eur. Respir. J.* 2018; 52.

9. Thulborn SJ, Mistry V, Brightling CE, Moffitt KL, Ribeiro D, Bafadhel M. Neutrophil elastase as a biomarker for bacterial infection in COPD. *Respir. Res.* 2019; 20: 1–7.
10. Wang Z, Locantore N, Haldar K, Ramsheh MY, Beech AS, Ma W, Brown JR, Tal-Singer R, Barer MR, Bafadhel M, Donaldson GC, Wedzicha JA, Singh D, Wilkinson TMA, Miller BE, Brightling CE. Inflammatory Endotype Associated Airway Microbiome in COPD Clinical Stability and Exacerbations – A Multi-Cohort Longitudinal Analysis. Am. J. Respir. Crit. Care Med. 2020. doi: 10.1164/rccm.202009-3448OC
11. Wiedermann FJ, Mayr AJ, Kaneider NC, Fuchs D, Mutz NJ, Schobersberger W. Alveolar Granulocyte Colony-Stimulating Factor and α-Chemokines in Relation to Serum Levels, Pulmonary Neutrophilia, and Severity of Lung Injury in ARDS. *Chest* 2004; 125: 212–219.
12. Wiedermann FJ, Mayr AJ, Hobisch-Hagen P, Fuchs D, Schobersberger W. Association of Endogenous G-CSF with Anti-Inflammatory Mediators in Patients with Acute Respiratory Distress Syndrome. *J. Interf. Cytokine Res.* 2003; 23: 729–736.
13. Aggarwal A, Baker CS, Evans TW, Haslam PL. G-CSF and IL-8 but not GM-CSF correlate with severity of pulmonary neutrophilia in acute respiratory distress syndrome. *Eur. Respir. J.* 2000; 15: 895–901.
14. Woodruff PG, Albert RK, Bailey WC, Casaburi R, Connett JE, Cooper JAD, Criner GJ, Curtis JL, Dransfield MT, Han MK, Harnden SM, Kim V, Marchetti N, Martinez FJ, McEvoy CE, Niewoehner DE, Reilly JJ, Rice K, Scanlon PD, Scharf SM, Sciurba FC, Washko GR, Lazarus SC. Randomized trial of zileuton for treatment of COPD exacerbations requiring hospitalization. *COPD J. Chronic Obstr. Pulm. Dis.* 2011; 8: 21–29.
15. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pẽa AG, Goodrich JK, Gordon JI, Huttley GA, Kelley ST, Knights D, Koenig JE, Ley RE, Lozupone CA, McDonald D, Muegge BD, Pirrung M, Reeder J, Sevinsky JR, Turnbaugh PJ, Walters WA, Widmann J, Yatsunenko T, Zaneveld J, Knight R. QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 2010; 7: 335–336.
16. Kopylova E, Navas-molina JA, Mercier C, Xu Z. Plate_11A.pdf. 2014; 1: 1–16.
17. McMurdie PJ, Holmes S. Waste Not, Want Not: Why Rarefying Microbiome Data Is

Inadmissible. *PLoS Comput. Biol.* 2014; 10.

18. Haynes W. Benjamini--Hochberg Method. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H, editors. *Encycl. Syst. Biol.* [Internet] New York, NY: Springer New York; 2013. p. 78Available from: https://doi.org/10.1007/978-1-4419-9863-7_1215.
19. Mayhew D, Devos N, Lambert C, Brown JR, Clarke SC, Kim VL, Magid-Slav M, Miller BE, Ostridge KK, Patel R, Sathe G, Simola DF, Staples KJ, Sung R, Tal-Singer R, Tuck AC, Van Horn S, Weynants V, Williams NP, Devaster JM, Wilkinson TMA. Longitudinal profiling of the lung microbiome in the AERIS study demonstrates repeatability of bacterial and eosinophilic COPD exacerbations. *Thorax* 2018; 73: 422–430.
20. Wang Z, Singh R, Miller BE, Tal-Singer R, Van Horn S, Tomsho L, MacKay A, Allinson JP, Webb AJ, Brookes AJ, George LM, Barker B, Kolsum U, Donnelly LE, Belchamber K, Barnes PJ, Singh D, Brightling CE, Donaldson GC, Wedzicha JA, Brown JR. Sputum microbiome temporal variability and dysbiosis in chronic obstructive pulmonary disease exacerbations: An analysis of the COPDMAP study. *Thorax* 2018; 73: 331–338.
21. Ghebre MA, Pang PH, Diver S, Desai D, Bafadhel M, Haldar K, Kebadze T, Cohen S, Newbold P, Rapley L, Woods J, Rugman P, Pavord ID, Johnston SL, Barer M, May RD, Brightling CE. Biological exacerbation clusters demonstrate asthma and chronic obstructive pulmonary disease overlap with distinct mediator and microbiome profiles. J. Allergy Clin. Immunol. 2018. p. 2027-2036.e12.
22. Wang Z, Bafadhel M, Haldar K, Spivak A, Mayhew D, Miller BE, Tal-Singer R, Johnston SL, Ramsheh MY, Barer MR, Brightling CE, Brown JR. Lung microbiome dynamics in COPD exacerbations. *Eur. Respir. J.* [Internet] 2016; 47: 1082–1092Available from: http://dx.doi.org/10.1183/13993003.01406-2015.
23. Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch S V. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *J. Clin. Microbiol.* 2014; 52: 2813–2823.
24. Keir HR, Dicker A, Lonergan M, Crichton M, Miller BE, Tal-Singer R, Chalmers JD. Clinical endotypes of exacerbation are associated with differences in microbial composition and diversity in COPD. *Eur. Respir. J.* 2020; : 2000391.
25. Dicker AJ, Crichton ML, Pumphrey EG, Cassidy AJ, Suarez-Cuartin G, Sibila O, Furrie E,

Fong CJ, Ibrahim W, Brady G, Einarsson GG, Elborn JS, Schembri S, Marshall SE, Palmer CNA, Chalmers JD. Neutrophil extracellular traps are associated with disease severity and microbiota diversity in patients with chronic obstructive pulmonary disease. J. Allergy Clin. Immunol. 2018. p. 117–127.

26. Bajrami B, Zhu H, Kwak HJ, Mondal S, Hou Q, Geng G, Karatepe K, Zhang YC, Nombela-Arrieta C, Park SY, Loison F, Sakai J, Xu Y, Silberstein LE, Luo HR. G-CSF maintains controlled neutrophil mobilization during acute inflammation by negatively regulating CXCR2 signaling. *J. Exp. Med.* 2016; 213: 1999–2018.
27. Semerad CL, Liu F, Gregory AD, Stumpf K, Link DC. G-CSF is an essential regulator of neutrophil trafficking from the bone marrow to the blood. *Immunity* 2002; 17: 413–423.
28. Hoenderdos K, Condliffe A. The neutrophil in chronic obstructive pulmonary disease: Too little, too late or too much, too soon? *Am. J. Respir. Cell Mol. Biol.* 2013; 48: 531–539.
29. Matute-Bello G, Lues WC, Radella F, Steinberg KP, Ruzinski JT, Jonas M, Chi EY, Hudson LD, Martin TR. Neutrophil apoptosis in the acute respiratory distress syndrome. *Pneumologie* 1998; 52: 354–355.
30. Van Raam BJ, Drewniak A, Groenewold V, Van Den Berg TK, Kuijpers TW. Granulocyte colony-stimulating factor delays neutrophil apoptosis by inhibition of calpains upstream of caspase-3. *Blood* 2008; 112: 2046–2054.
31. Salzer AT, Niemiec MJ, Hosseinzadeh A, Stylianou M, Åström F, Röhm M, Ahlm C, Wahlin A, Ermert D, Urban CF. Assessment of neutrophil chemotaxis upon G-CSF treatment of healthy stem cell donors and in allogeneic transplant recipients. *Front. Immunol.* 2018; 9: 1–11.
32. Castellani S, D'Oria S, Diana A, Polizzi AM, Di Gioia S, Mariggiò MA, Guerra L, Favia M, Vinella A, Leonetti G, De Venuto D, Gallo C, Montemurro P, Conese M. G-CSF and GM-CSF Modify Neutrophil Functions at Concentrations found in Cystic Fibrosis. *Sci. Rep.* 2019; 9: 1–11.

Table 1. Patient characteristics

	LEUKO cohort		MRC cohort
	Placebo	Zileuton	
Subjects (n)	33	36	59
Age (years)	63 ± 1.7	61 ± 1.5	70 ± 9.3
Gender (n/%)	Male = 22 (67) Female = 11 (33)	Male = 20 (56) Female = 16 (44)	Male = 41 (70) Female = 18 (30)
Race (n/%)	Caucasian = 14 (42) African-American =18 (54) Hispanic = 2 (6) Native-American = 3 (9)	Caucasian = 21 (58) African-American = 13 (36) Hispanic = 3 (8) Native-American = 4 (11)	Caucasian = 58 (98) Black = 1 (2)
Smoking status (n/%)	Ex-smoker = 21 (64) Current smoker = 12 (36)	Ex-smoker = 25 (70); Current = 11 (30)	Ex-smoker = 41 (70) Current = 15 (25) Non-smoker = 3 (5)
GOLD stage (n/%)			GOLD 2 = 28 (47) GOLD 3 = 22 (37) GOLD 4 = 9 (15)
pre-bronchodilator FEV1 (Mean ± SEM)	0.87 ± 0.09	0.75 ± 0.17	1.31 ± 0.07
post-bronchodilator FEV1 (Mean ± SEM)	0.81 ± 0.11	1.0 ± 0.18	1.36 ± 0.07
FVC (Mean ± SEM)	2.46 ± 0.96	2.59 ± 0.90	2.67 ± 0.09 missing 1 subject
pre-bronchodilator FEV/FVC (Mean ± SEM)	0.39 ± 0.04	0.44 ± 0.04	0.49 ± 0.02
post-bronchodilator FEV/FVC (Mean ± SEM)	0.47 ± 0.05	0.45 ± 0.04	0.51 ± 0.02

*37 LEUKO participants with matched G-CSF measurements of Exacerbation (D0) and Recovery (D30)

FIGURE LEGENDS

Figure 1. Serum G-CSF is significantly elevated during COPD exacerbation A) Serum G-CSF was measured at exacerbation (N=54) and recovery (30-day post exacerbation, N=52) from a hospitalized AECOPD cohort (LEUKO cohort). Paired exacerbation and recovery samples were available for 37 patients, and analysis was performed on pooled treatment arms as there was no obvious treatment advantage of zileuton compared with placebo. P=Wilcoxon matched-pairs signed rank test. B) Among 37 paired samples, 21 subjects had $\geq$ 2-fold decrease of serum G-CSF at recovery. Medians $\pm$ 95% CI and Mann-Whitney U test p values shown. C) A heatmap of Spearman's correlations between the fold change of the biomarkers at exacerbation compared to a preceding stable visit. Each square represents the correlation between the feature heading the column with the feature heading the row. The number in a given square indicates the corresponding Spearman correlation coefficient between the two features.

Figure 2. Elevated levels of serum G-CSF during exacerbations are non-eosinophilic and predominantly associated with infection. A) Serum G-CSF was measured from 74 exacerbations with preceding stable visits and at 2 weeks and 6 weeks follow-up visits from 54 subjects (MRC cohort). B-C) Serum G-CSF measurements were categorized between eosinophilic and non-eosinophilic exacerbations. Serum G-CSF measurements were compared between exacerbations strictly associated with D) bacterial infection (N=18), E) viral infection (N=7) or F) infectious defined by bacteria, virus or bacteria/virus co-infection (N=37). Statistical significance is shown by respective p-values generated by Wilcoxon matched-pairs signed rank test.

Figure 3. Exacerbations categorized by G-CSF induction are characterized by higher biomarkers of neutrophilic inflammation. G-CSF high (Hi) and low (Lo) groups are categorized based on $\geq$2-fold higher G-CSF at exacerbation compared to preceding stable visit. A) Blood neutrophil fold change is the change of blood neutrophil at exacerbation compared to preceding stable visit, B) absolute blood neutrophil count at exacerbation, C) serum IL-6 at exacerbation, D) serum SAA at exacerbation, E) sputum neutrophil percentage at exacerbation, and F) sputum IL-6 at exacerbation. Mann-Whitney U test p values are shown. G) A heatmap of Spearman's correlations between biomarkers at exacerbation. H) A heatmap of Spearman's correlations

between the fold change of the biomarkers at exacerbation compared to the preceding stable visit. Each square represents the correlation between the feature heading the column with the feature heading the row. The number in a given square indicates the corresponding Spearman correlation coefficient between the two features.

Figure 4. High G-CSF levels at exacerbation is associated with higher frequency of bacterial infection. A-C,E) Biomarker levels were compared between bacterial vs non-bacterial exacerbations. A) G-CSF fold change between exacerbation and preceding stable visit, B) serum G-CSF at exacerbation, C) sputum IL-1β at exacerbation, E) serum CRP at exacerbation. D,F) Spearman's correlation plots are shown between serum G-CSF at exacerbation and D) sputum IL-1β at exacerbation, and F) serum CRP at exacerbation. G) Receiver operating characteristic curve illustrating biomarkers that positively predict bacterial exacerbation, AUC=area under the curve. A-C,E) Medians ± 95% CI and Mann-Whitney U test p values are shown. D,F) Spearman's correlations and p values are indicated.

Figure 5. An altered sputum microbiome precedes exacerbations characterized by bacteria or ≥2 fold increase in serum G-CSF. A,E) Shannon's Diversity Index, B,F) NMDS ordination plot of Bray-Curtis distances with dashed ellipses representing the 95% CI for the centroid of each stratification group, C,G) *Haemophilus* abundance, and D,H) *Moraxella* abundance for the sputum microbiota during stable disease prior to bacterial exacerbations. A-D) Bacterial exacerbations, No: n=17, Yes: n=22. E-H) Exacerbations characterized by ≥2-fold increase in G-CSF, Lo: n=33, Hi: n=13). For a and e, group mean +/- standard error and Wilcoxon Rank-Sum test p values are shown. For B and F, permutational ANOVA was calculated. For C, D, G, and H, the y-axes have been fixed to the limits of log10(relative abundance) for these samples, group mean log10(relative abundance) +/- standard error is depicted, and statistical significance was determined by DESeq2 and adjusted for false discoveries.

Figure 6. Serum G-CSF levels at exacerbation stratifies neutrophilic-dysbiotic exacerbations. A) serum G-CSF levels, B) blood neutrophil counts, C) serum CRP levels at exacerbation were compared between Neutrophilic (sputum neutrophil ≥61%) and Neutrophil-Lo COPD

exacerbations. D) Serum G-CSF levels, E) blood neutrophil counts, F) serum CRP levels at exacerbation were compared between Neutrophilic-dysbiotic (relative abundance of Haemophilus and Moraxella >0.41), Neutrophilic-balanced and Neutrophil-Lo (sputum neutrophil <61%). G) Receiver operating characteristic curve for classifying Neutrophilic-dysbiotic COPD exacerbations, AUC=area under the curve. Mann-Whitney U test p values are shown.

Figure 7. IL-1β treatment directly induces genes conducive of neutrophilic inflammation. A-C) Normal human bronchial epithelial cells were differentiated to air-liquid interface and treated with IL-1β (10ng/ml) for 24 h and the cytokine levels were measured from cell supernatants by Luminex based ELISA assay D-H) RNA was isolated to examine the gene expression by RNAseq. Statistical significance was determined by unpaired t-test with Welch's correction.

Supplementary Figures

Supplementary figure 1.

A) Spearman's correlation of G-CSF levels at exacerbation with G-CSF fold change during recovery (N=37). B) Subjects having high G-CSF levels at exacerbation has shorter length of hospital stay following antibiotic (Abx) treatment as represented by the Box-plot and C) Kaplan-Meier plot.

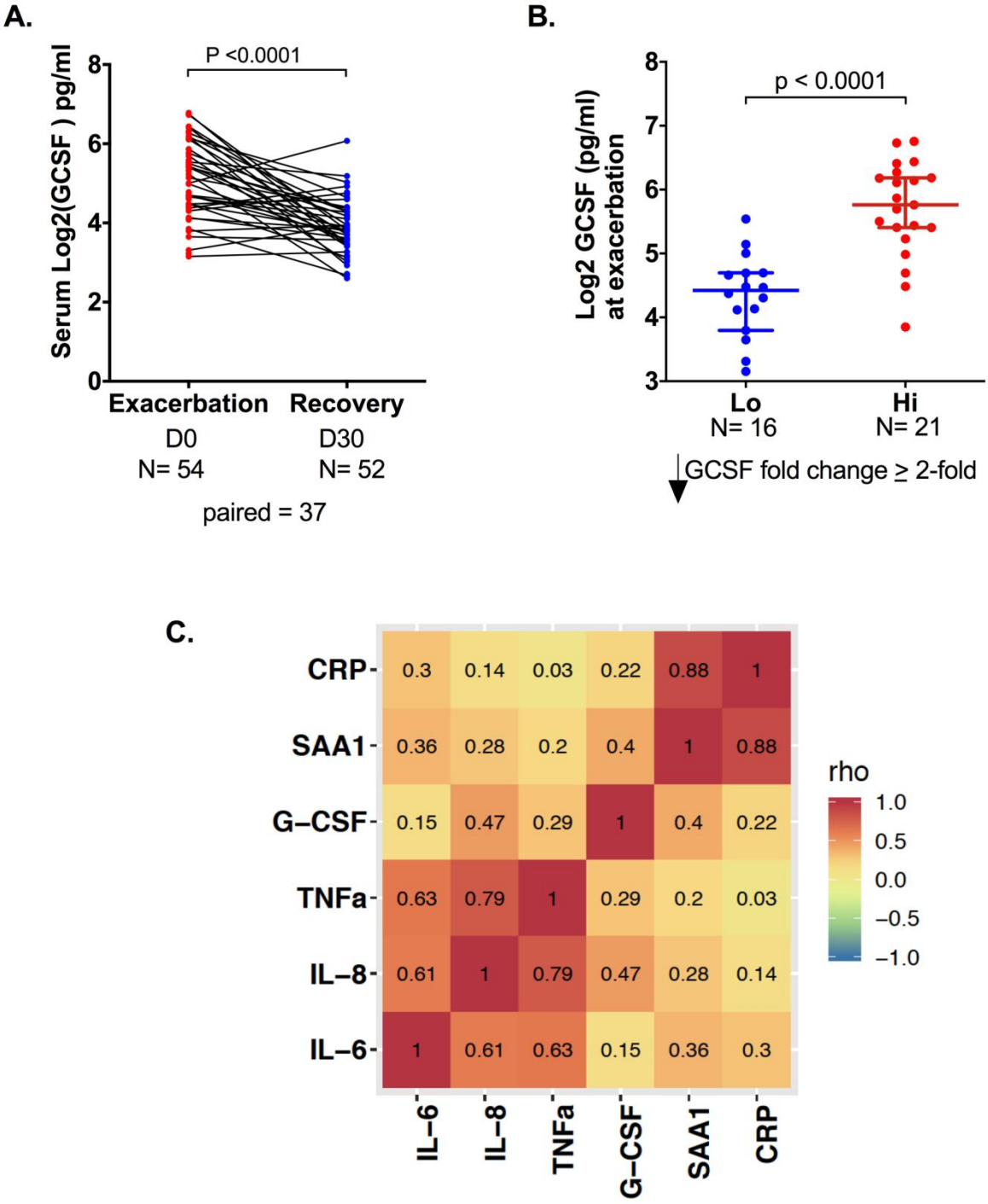
Supplementary figure 2. High serum G-CSF during exacerbation reflects lung neutrophilic inflammation. A-B) G-CSF fold change (N=68) and G-CSF levels at exacerbation (N=68) were plotted as a continuous variable against sputum neutrophil percent at exacerbation. C-D) G-CSF fold change (N=62) and G-CSF levels at exacerbation (N=62) were plotted as a continuous variable against sputum IL-6 levels at exacerbation. Statistical significance analyzed by Spearman's correlation.

Supplementary figure 3.

Baseline G-CSF is prognostic of neutrophilic inflammation during exacerbation. Patients were grouped into 4 quartiles based on their baseline G-CSF expression levels and plotted against A) blood neutrophil counts, B) sputum neutrophil percent, C) sputum IL-8 and D) serum G-CSF at exacerbation. Wilcoxon rank sum test p values were calculated between the lowest and the highest quartiles.

Supplementary figure 4. An altered sputum microbiome is present during both bacterial exacerbations and those with ≥2- fold increase in G-CSF. A,E) Shannon's Diversity Index, B,F) NMDS ordination plot of Bray-Curtis distances with dashed ellipses representing the 95% CI for the centroid of each stratification group, C,G) *Haemophilus* abundance, and D,H) *Moraxella* abundance for the sputum microbiota during bacterial exacerbations. A-D) Bacterial exacerbations, No: n=16, Yes: n=17. E-H) exacerbations characterized by >2-fold increase in G-CSF, No: n=31, Yes: n=9). For A and E, group mean +/- standard error and Wilcoxon Rank-Sum test p values are shown. For B and F, permutational ANOVA was calculated. For C, D, G, and H, the y-axes have been fixed to the limits of log10(relative abundance) for these samples, group mean log10(relative abundance) +/- standard error is depicted, and statistical significance was determined by DESeq2 and adjusted for false discoveries.

Supplementary figure 5. Model summary. A dysbiotic lung microbiota with lower bacterial diversity driven by colonization of potential respiratory pathogens such as *Haemophilus* and *Moraxella* precedes bacterial exacerbations in COPD patients. Elevated serum G-CSF at exacerbation classifies exacerbations characterized by bacteria (increased abundance of respiratory pathogens such as *Haemophilus* and *Moraxella,* increased bacterial load, and reduced lung microbiome diversity) and neutrophilic inflammation (high sputum and blood neutrophils and sputum IL-1β).


A.
P <0.0001
Serum Log2(GCSF) pg/ml
8
6
4
2
0
Exacerbation
Recovery
D0
D30
N= 54
N= 52
paired = 37
B.
p < 0.0001
Log2 GCSF (pg/ml) at exacerbation
8
7
6
5
4
3
Lo
Hi
N= 16
N= 21
GCSF fold change ≥ 2-fold
C.
CRP
SAA1
G–CSF
TNFa
IL–8
IL–6
0.3 0.14 0.03 0.22 0.88 1
0.36 0.28 0.2 0.4 1 0.88
0.15 0.47 0.29 1 0.4 0.22
0.63 0.79 1 0.29 0.2 0.03
0.61 1 0.79 0.47 0.28 0.14
1 0.61 0.63 0.15 0.36 0.3
IL–6
IL–8
TNFa
G–CSF
SAA1
CRP
rho
1.0
0.5
0.0
−0.5
−1.0

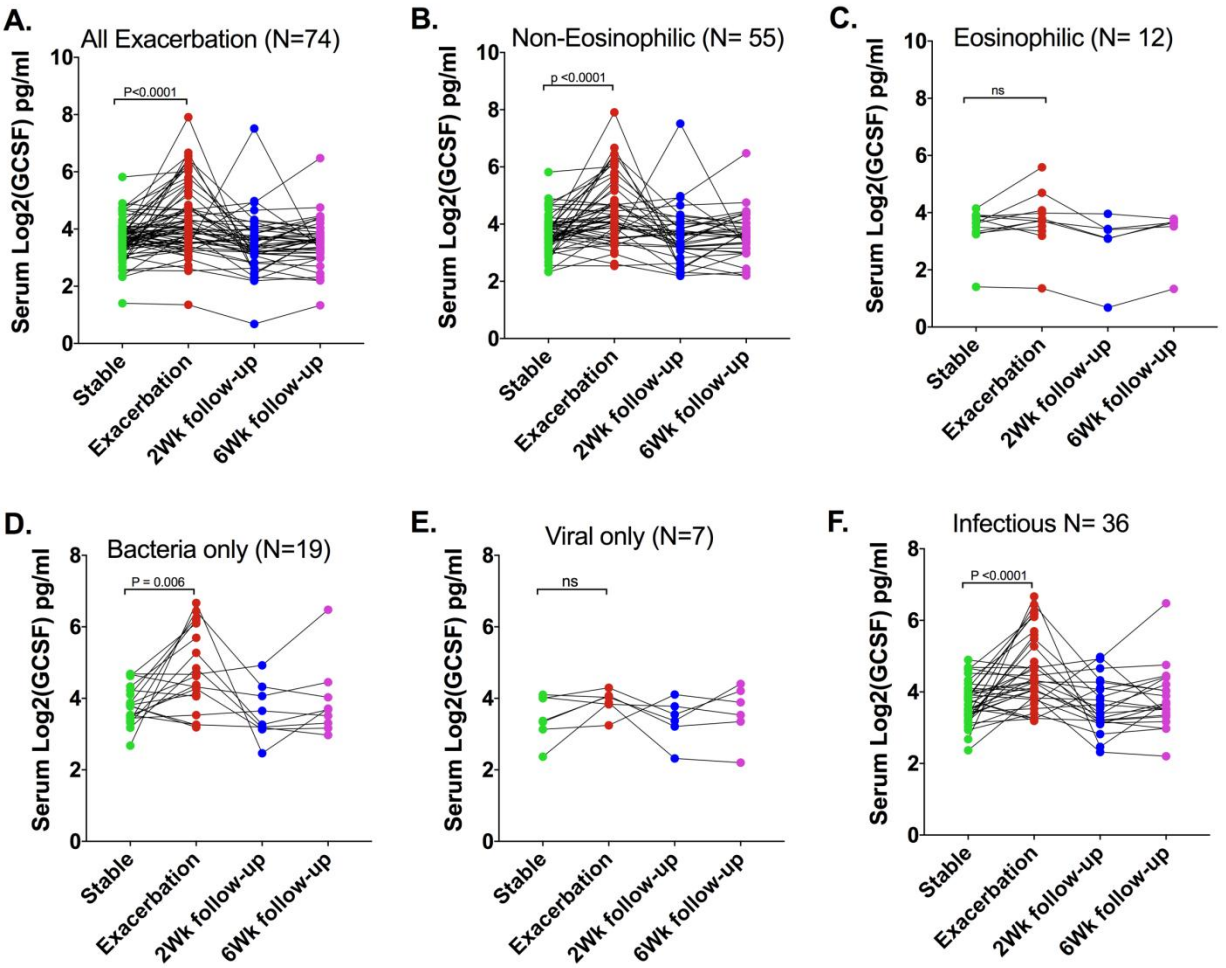

A.
All Exacerbation (N=74)
P<0.0001
Serum Log2(GCSF) pg/ml
B.
Non-Eosinophilic (N= 55)
p <0.0001
C.
Eosinophilic (N= 12)
ns
D.
Bacteria only (N=19)
P = 0.006
E.
Viral only (N=7)
ns
F.
Infectious N= 36
P <0.0001
Stable
Exacerbation
2Wk follow-up
6Wk follow-up

Figure 2

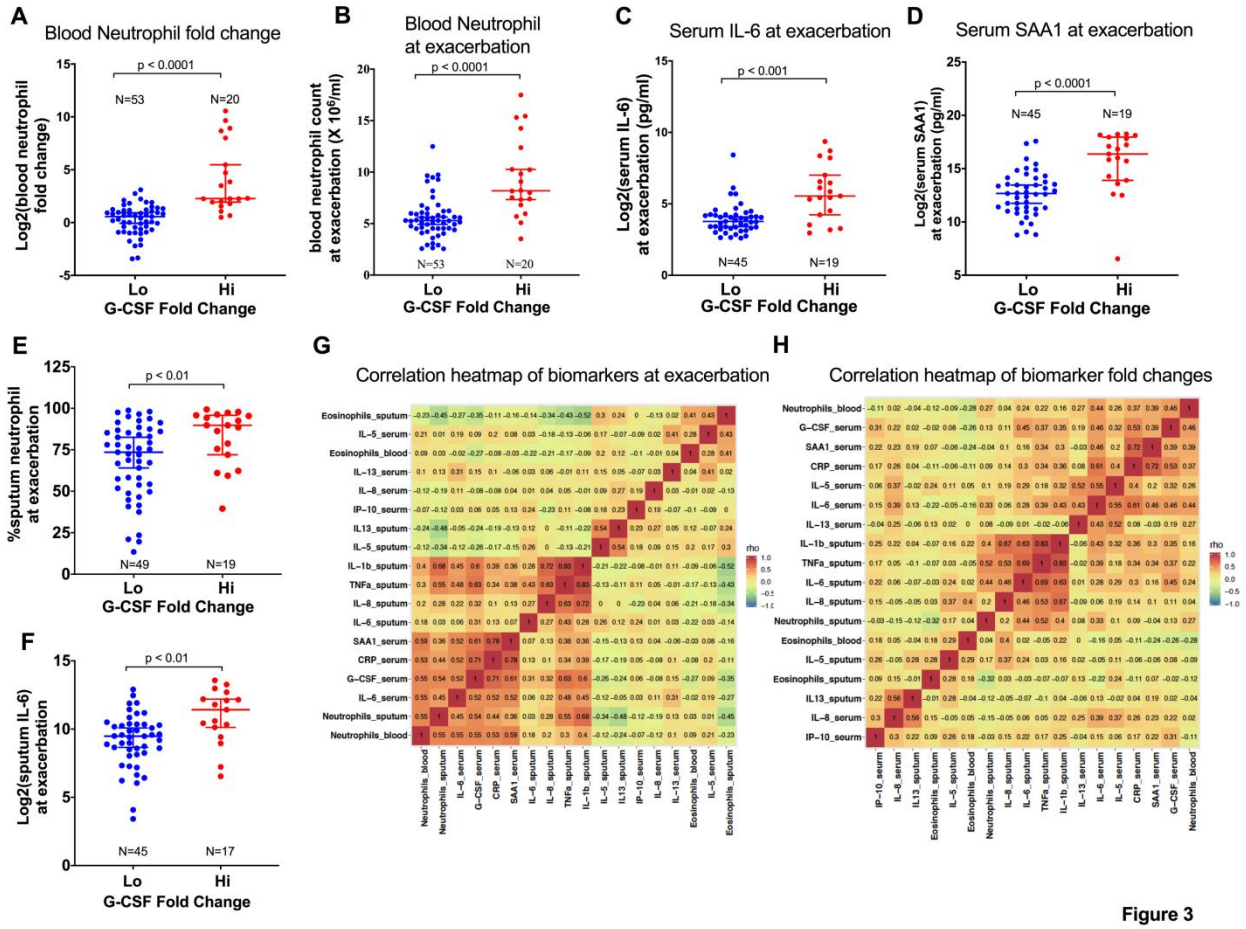

A
Blood Neutrophil fold change
Log2(blood neutrophil fold change)
p < 0.0001
N=53
N=20
Lo
Hi
G-CSF Fold Change
B
Blood Neutrophil at exacerbation
blood neutrophil count at exacerbation (X 10^6/ml)
p < 0.0001
N=53
N=20
Lo
Hi
G-CSF Fold Change
C
Serum IL-6 at exacerbation
Log2(serum IL-6) at exacerbation (pg/ml)
p < 0.001
N=45
N=19
Lo
Hi
G-CSF Fold Change
D
Serum SAA1 at exacerbation
Log2(serum SAA1) at exacerbation (pg/ml)
p < 0.0001
N=45
N=19
Lo
Hi
G-CSF Fold Change
E
%sputum neutrophil at exacerbation
p < 0.01
N=49
N=19
Lo
Hi
G-CSF Fold Change
F
Log2(sputum IL-6) at exacerbation
p < 0.01
N=45
N=17
Lo
Hi
G-CSF Fold Change
G
Correlation heatmap of biomarkers at exacerbation
Eosinophils_sputum
IL-5_serum
Eosinophils_blood
IL-13_serum
IL-8_serum
IP-10_serum
IL13_sputum
IL-5_sputum
IL-1b_sputum
TNFa_sputum
IL-8_sputum
IL-6_sputum
SAA1_serum
CRP_serum
G-CSF_serum
IL-6_serum
Neutrophils_sputum
Neutrophils_blood
rho
H
Correlation heatmap of biomarker fold changes
Neutrophils_blood
G-CSF_serum
SAA1_serum
CRP_serum
IL-5_serum
IL-6_serum
IL-13_serum
IL-1b_sputum
TNFa_sputum
IL-6_sputum
IL-8_sputum
Neutrophils_sputum
Eosinophils_blood
IL-5_sputum
Eosinophils_sputum
IL13_sputum
IL-8_serum
IP-10_serum
rho

Figure 3

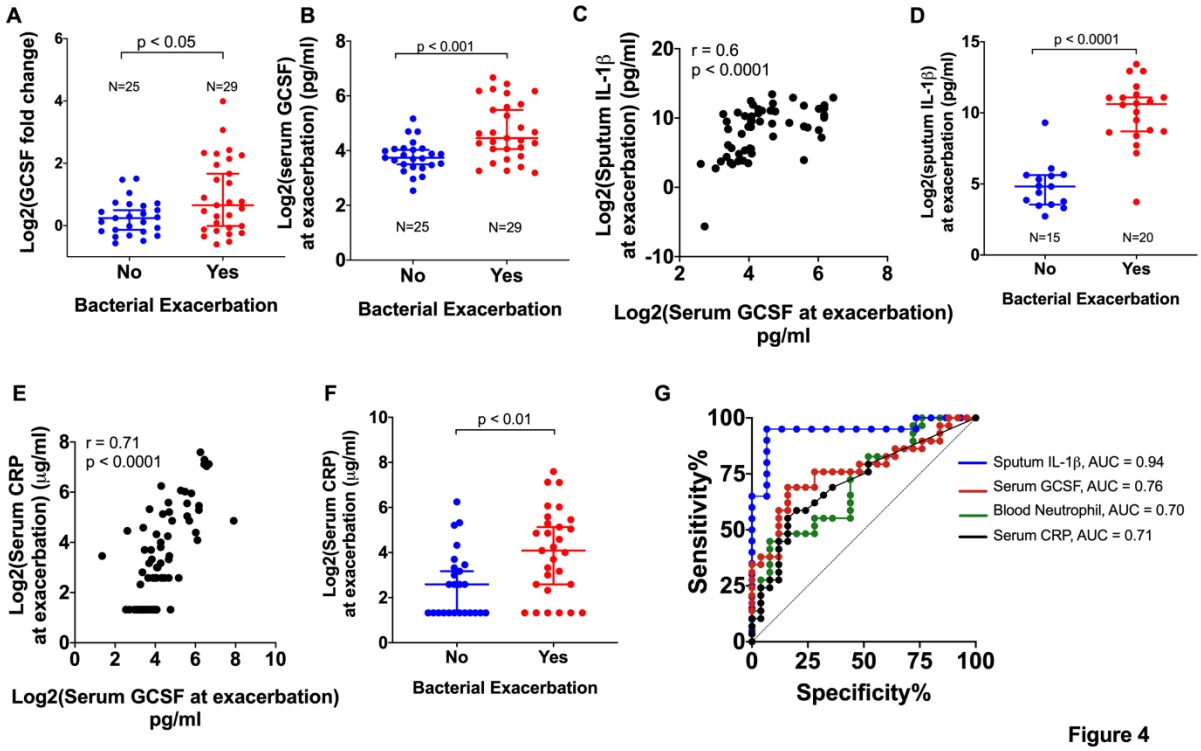
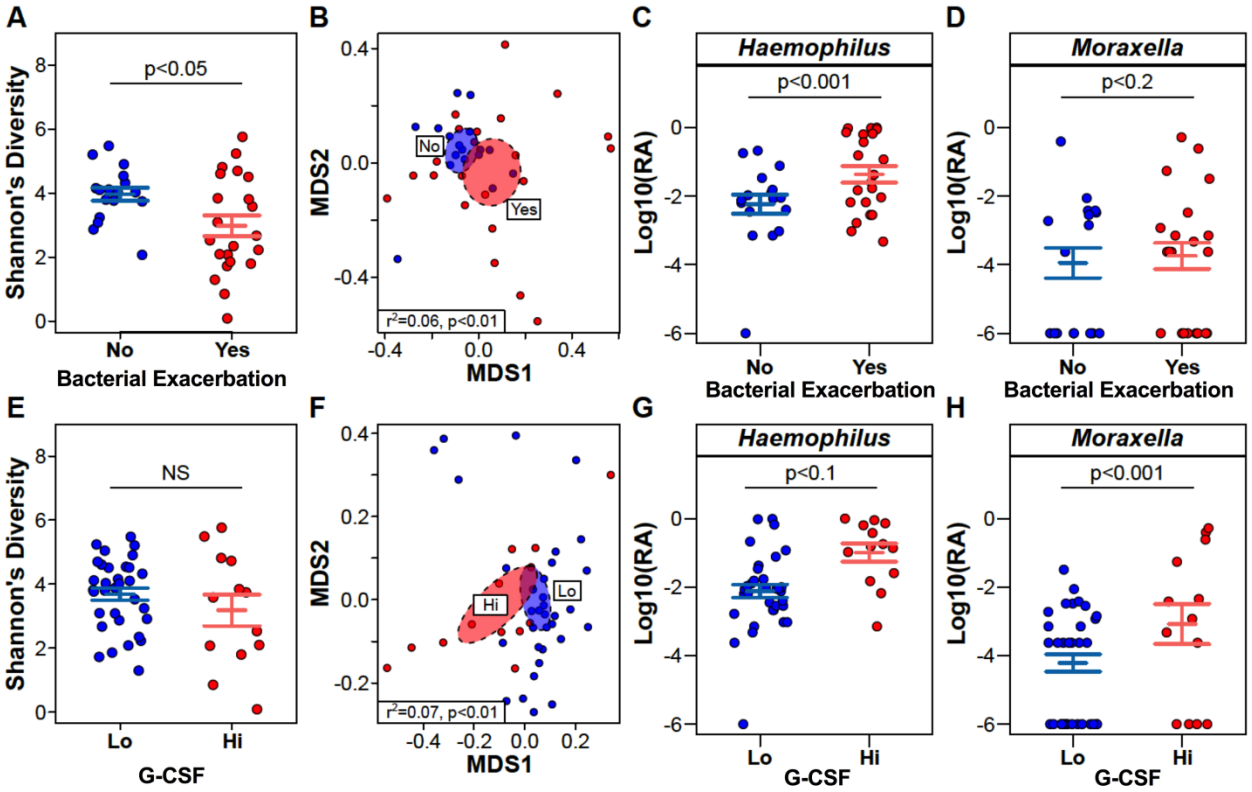

A
p < 0.05
N=25
N=29
Log2(GCSF fold change)
No
Yes
Bacterial Exacerbation
B
p < 0.001
Log2(serum GCSF) at exacerbation) (pg/ml)
N=25
N=29
No
Yes
Bacterial Exacerbation
C
r = 0.6
p < 0.0001
Log2(Sputum IL-1β at exacerbation) (pg/ml)
Log2(Serum GCSF at exacerbation) pg/ml
D
p < 0.0001
Log2(sputum IL-1β) at exacerbation (pg/ml)
N=15
N=20
No
Yes
Bacterial Exacerbation
E
r = 0.71
p < 0.0001
Log2(Serum CRP at exacerbation) (µg/ml)
Log2(Serum GCSF at exacerbation) pg/ml
F
p < 0.01
Log2(Serum CRP) at exacerbation (µg/ml)
No
Yes
Bacterial Exacerbation
G
Sensitivity%
Specificity%
Sputum IL-1β, AUC = 0.94
Serum GCSF, AUC = 0.76
Blood Neutrophil, AUC = 0.70
Serum CRP, AUC = 0.71

Figure 4


A
Shannon's Diversity
p<0.05
No
Yes
Bacterial Exacerbation
B
MDS2
MDS1
No
Yes
r²=0.06, p<0.01
C
Haemophilus
p<0.001
Log10(RA)
No
Yes
Bacterial Exacerbation
D
Moraxella
p<0.2
Log10(RA)
No
Yes
Bacterial Exacerbation
E
Shannon's Diversity
NS
Lo
Hi
G-CSF
F
MDS2
MDS1
Hi
Lo
r²=0.07, p<0.01
G
Haemophilus
p<0.1
Log10(RA)
Lo
Hi
G-CSF
H
Moraxella
p<0.001
Log10(RA)
Lo
Hi
G-CSF

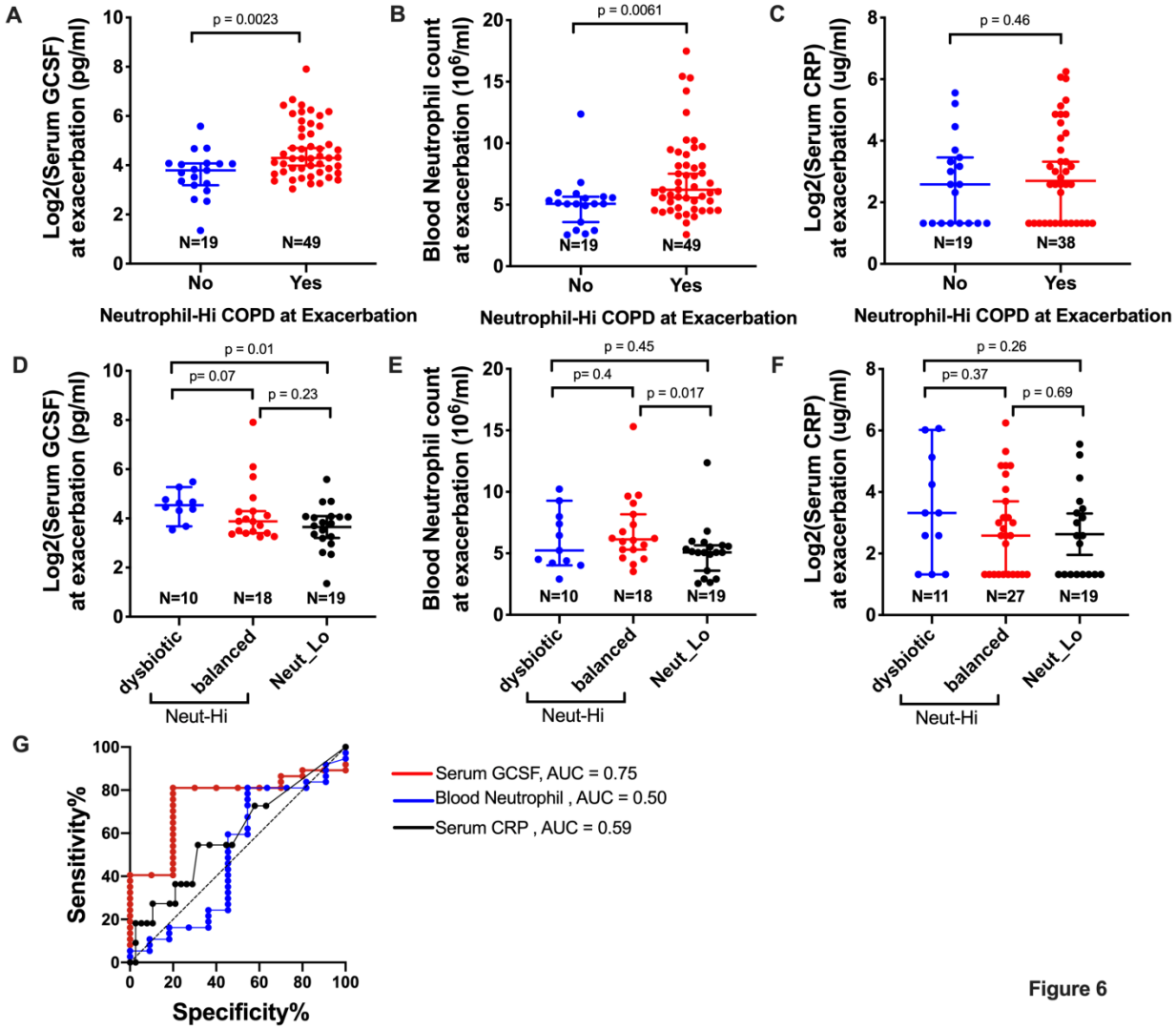

A
Log2(Serum GCSF) at exacerbation (pg/ml)
p = 0.0023
N=19
N=49
No
Yes
Neutrophil-Hi COPD at Exacerbation
B
Blood Neutrophil count at exacerbation (10^6/ml)
p = 0.0061
N=19
N=49
No
Yes
Neutrophil-Hi COPD at Exacerbation
C
Log2(Serum CRP) at exacerbation (ug/ml)
p = 0.46
N=19
N=38
No
Yes
Neutrophil-Hi COPD at Exacerbation
D
Log2(Serum GCSF) at exacerbation (pg/ml)
p = 0.01
p= 0.07
p = 0.23
N=10
N=18
N=19
dysbiotic
balanced
Neut_Lo
Neut-Hi
E
Blood Neutrophil count at exacerbation (10^6/ml)
p = 0.45
p= 0.4
p = 0.017
N=10
N=18
N=19
dysbiotic
balanced
Neut_Lo
Neut-Hi
F
Log2(Serum CRP) at exacerbation (ug/ml)
p = 0.26
p= 0.37
p = 0.69
N=11
N=27
N=19
dysbiotic
balanced
Neut_Lo
Neut-Hi
G
Sensitivity%
Specificity%
Serum GCSF, AUC = 0.75
Blood Neutrophil , AUC = 0.50
Serum CRP , AUC = 0.59

Figure 6

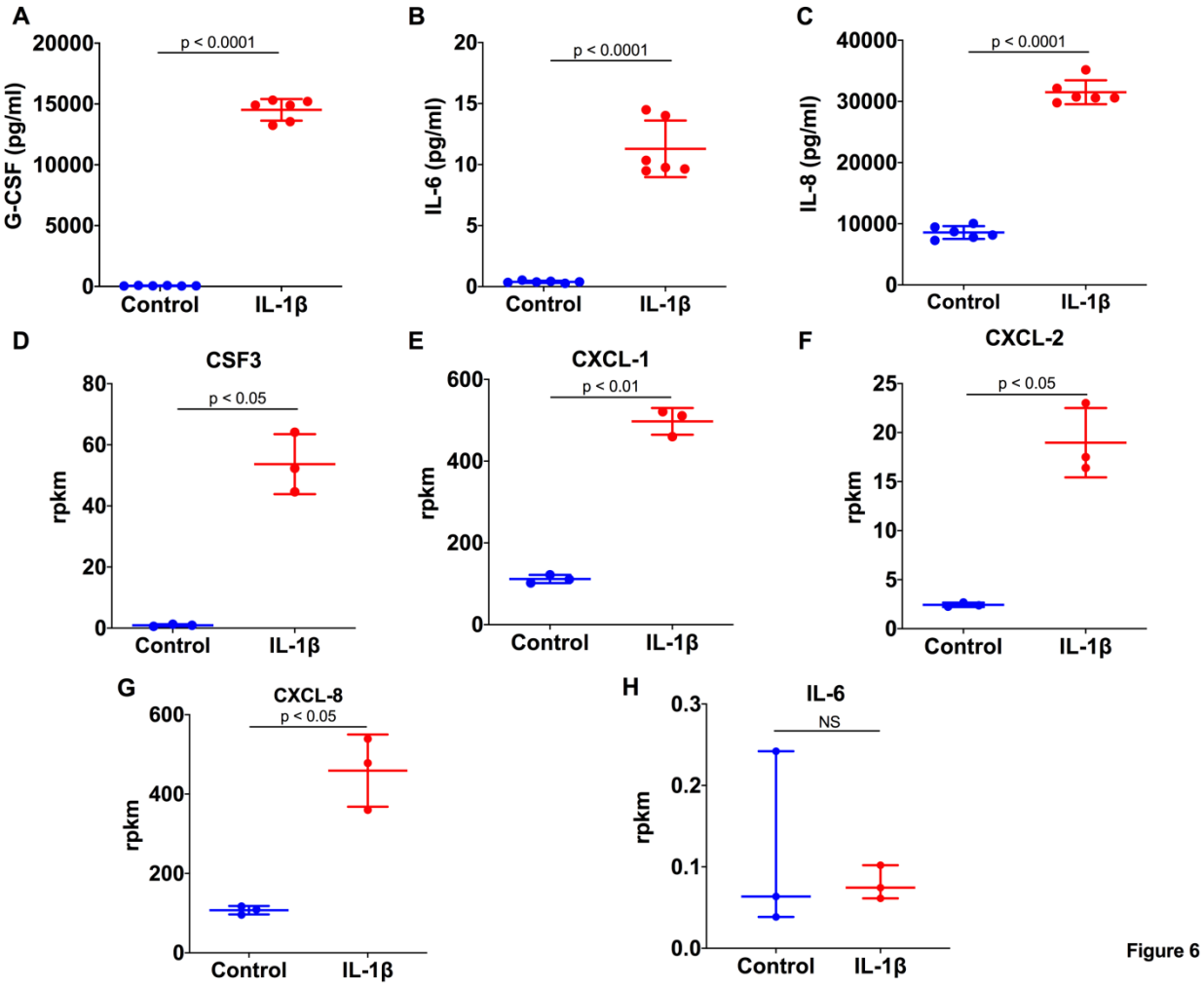
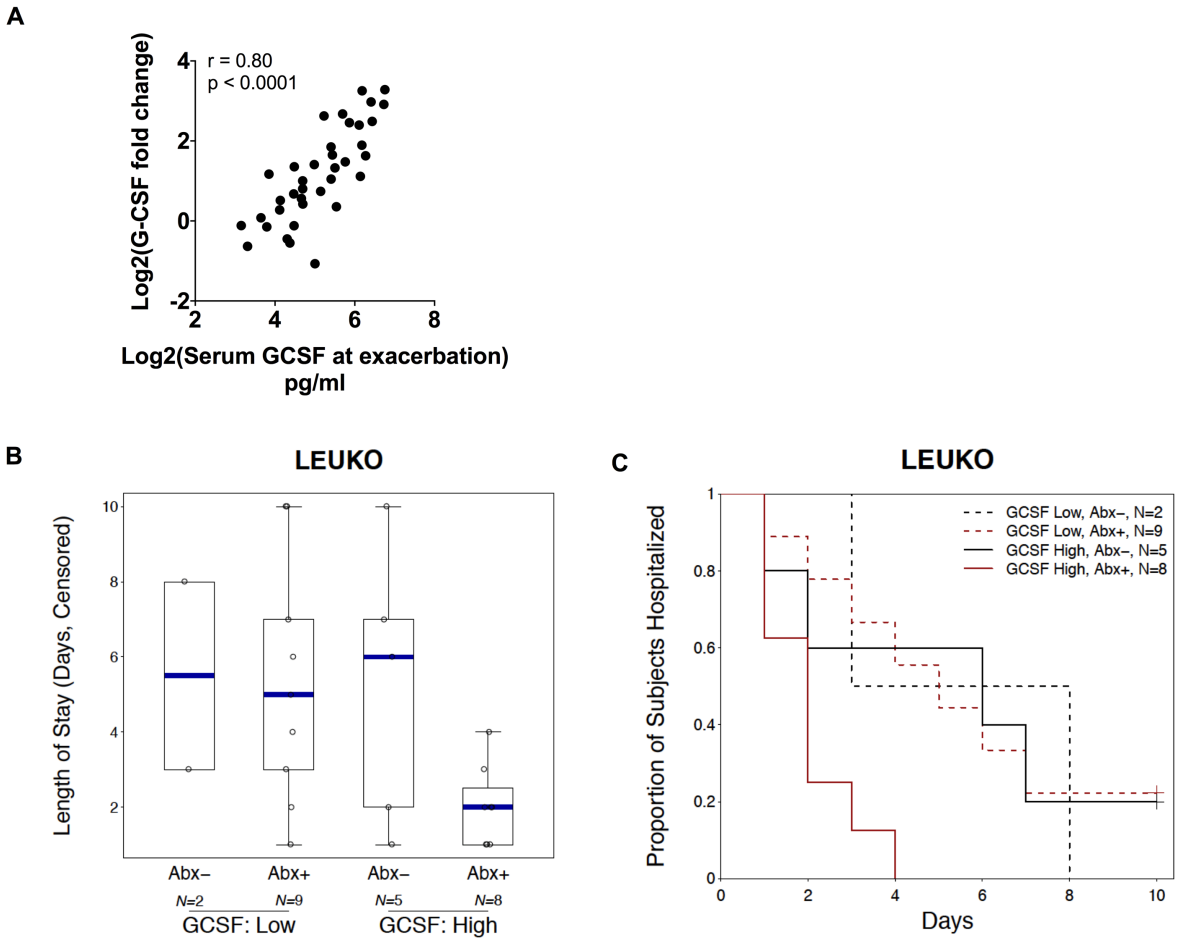

A
G-CSF (pg/ml)
20000
15000
10000
5000
0
p < 0.0001
Control
IL-1β
B
IL-6 (pg/ml)
20
15
10
5
0
p < 0.0001
Control
IL-1β
C
IL-8 (pg/ml)
40000
30000
20000
10000
0
p < 0.0001
Control
IL-1β
D
CSF3
rpkm
80
60
40
20
0
p < 0.05
Control
IL-1β
E
CXCL-1
rpkm
600
400
200
0
p < 0.01
Control
IL-1β
F
CXCL-2
rpkm
25
20
15
10
5
0
p < 0.05
Control
IL-1β
G
CXCL-8
rpkm
600
400
200
0
p < 0.05
Control
IL-1β
H
IL-6
rpkm
0.3
0.2
0.1
0.0
NS
Control
IL-1β

Figure 6

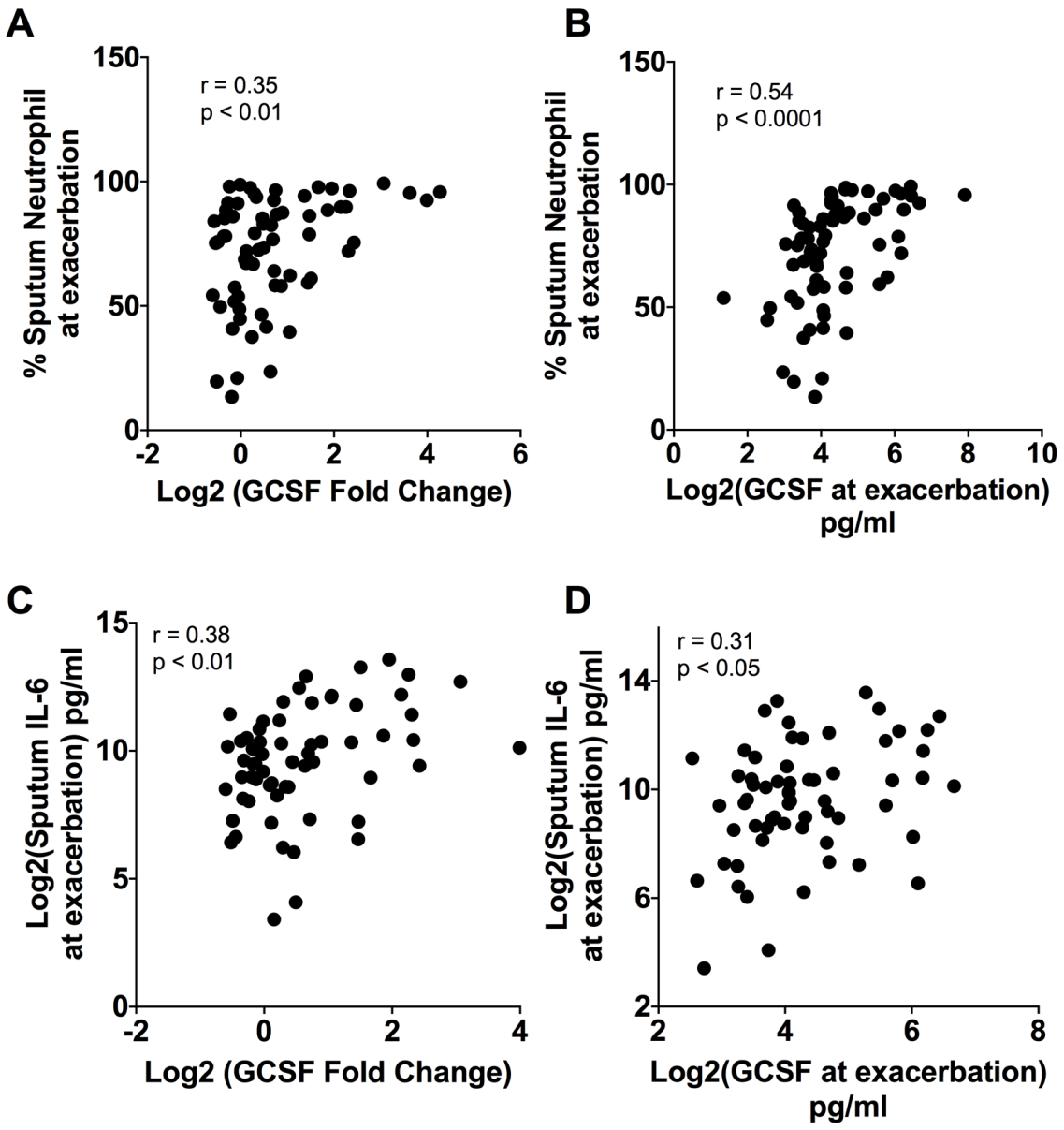
Supplementary Data

Supplementary Table 1. Subtypes of primary exacerbations (MRC cohort)

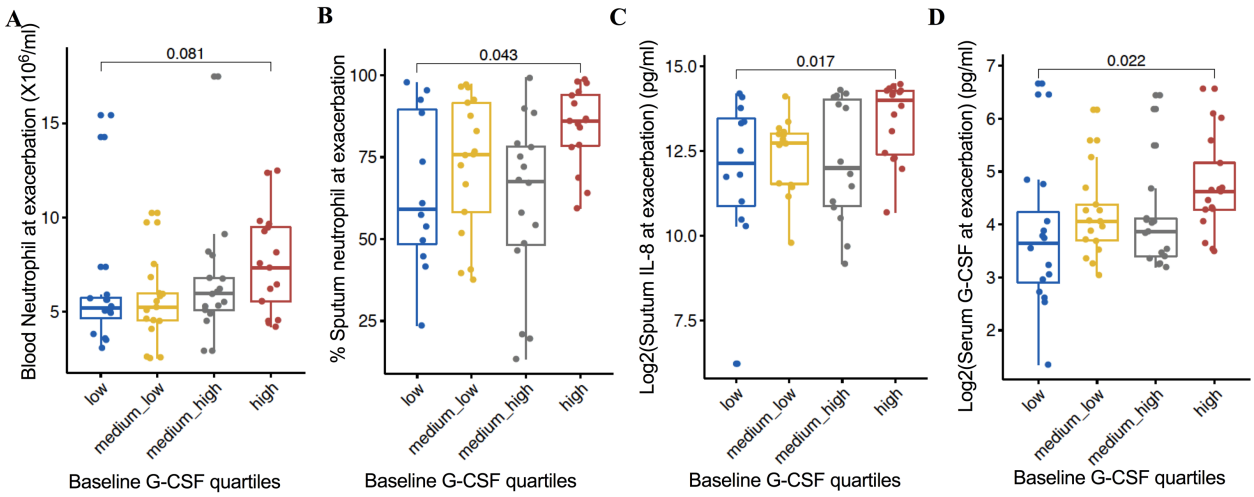
Exacerbation type	Number of primary exacerbations	Percent	Number of subjects
All	74	100%	59
Bacterial exacerbation	18	24.3%	18
Bacteria : Virus	9	12.2%	9
Bacteria: Eos	1	1.4%	1
Bacteria : Virus: Eos	1	1.4%	1
Viral Exacerbation	7	9.5%	7
Virus: Eos	1	1.4%	1
Eosinophilic Exacerbation	9	15.3%	9
Total infection associated exacerbation	37	50%	35



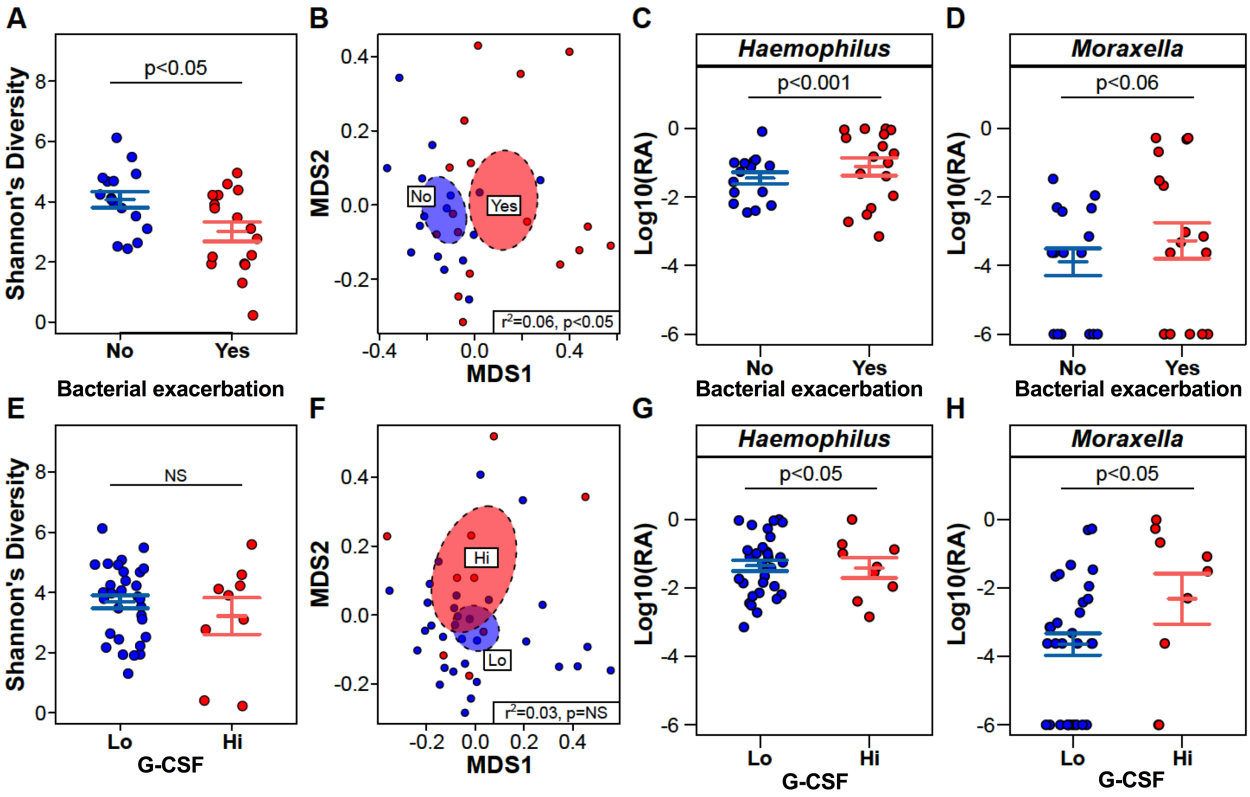
Supplementary figure 1. A) Spearman's correlation of G-CSF levels at exacerbation with G-CSF fold change during recovery (N=37). B) Subjects having high G-CSF levels at exacerbation has shorter length of hospital stay following antibiotic (Abx) treatment as represented by the Box-plot and C) Kaplan-Meier plot.



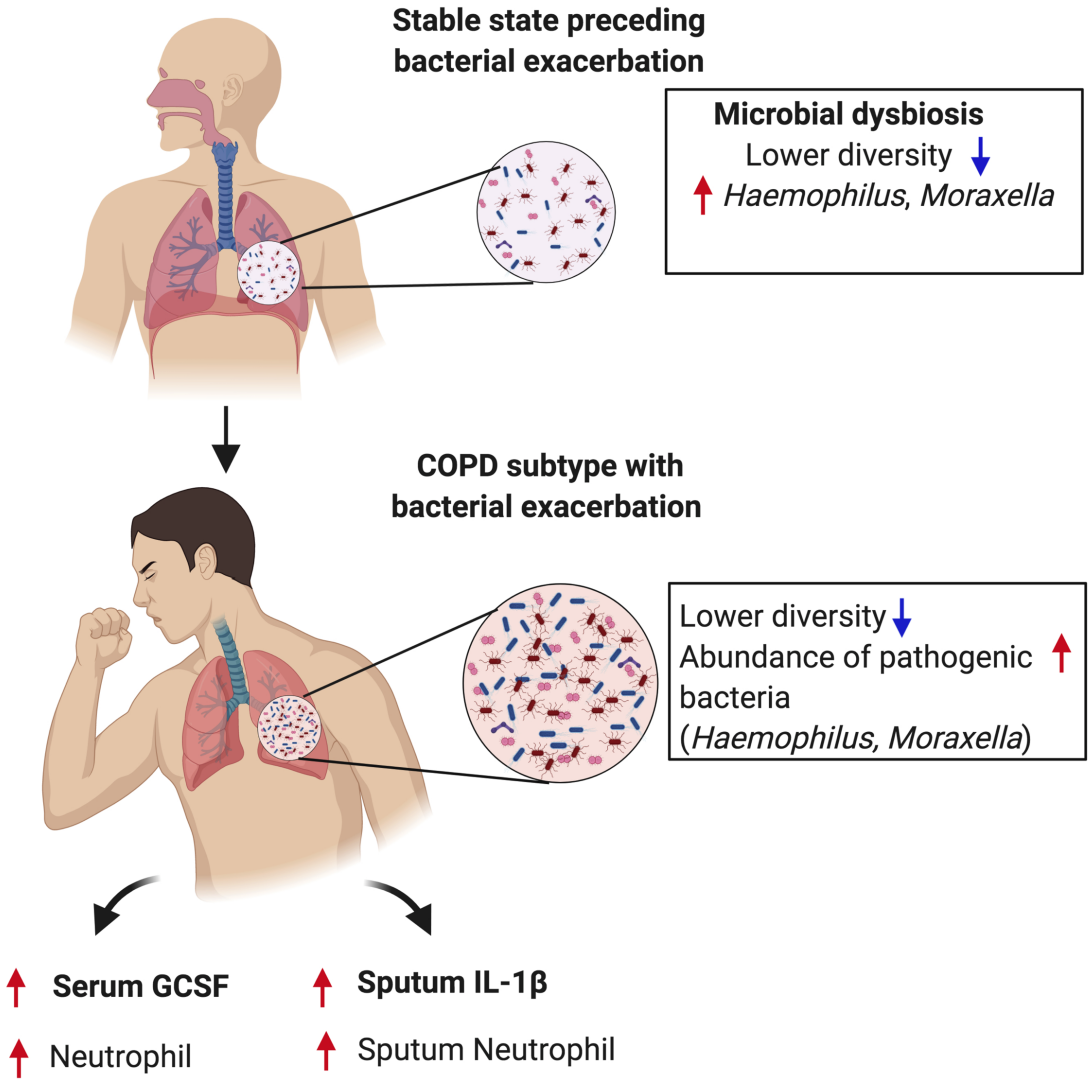
Supplementary figure 2. High serum G-CSF during exacerbation reflects lung neutrophilic inflammation. A-B) G-CSF fold change (N=68) and G-CSF levels at exacerbation (N=68) were plotted as a continuous variable against sputum neutrophil percent at exacerbation. C-D) G-CSF fold change (N=62) and G-CSF levels at exacerbation (N=62) were plotted as a continuous variable against sputum IL-6 levels at exacerbation. Statistical significance analyzed by Spearman's correlation.



Supplementary figure 3. Baseline G-CSF is prognostic of neutrophilic inflammation during exacerbation. Patients were grouped into 4 quartiles based on their baseline G-CSF expression levels and plotted against A) blood neutrophil counts, B) sputum neutrophil percent, C) sputum IL-8 and D) serum G-CSF at exacerbation. Wilcoxon rank sum test p values were calculated between the lowest and the highest quartiles.



Supplementary figure 4. An altered sputum microbiome is present during both bacterial exacerbations and those with ≥2- fold increase in G-CSF. A,E) Shannon's Diversity Index, B,F) NMDS ordination plot of Bray-Curtis distances with dashed ellipses representing the 95% CI for the centroid of each stratification group, C,G) *Haemophilus* abundance, and D,H) *Moraxella* abundance for the sputum microbiota during bacterial exacerbations. A-D) Bacterial exacerbations, No: n=16, Yes: n=17. E-H) exacerbations characterized by >2 fold increase in G-CSF, No: n=31, Yes: n=9). For A and E, group mean +/- standard error and Wilcoxon Rank-Sum test p values are shown. For B and F, permutational ANOVA was calculated. For C, D, G, and H, the y-axes have been fixed to the limits of log10(relative abundance) for these samples, group mean log10(relative abundance) +/- standard error is depicted, and statistical significance was determined by DESeq2 and adjusted for false discoveries.



Supplementary figure 5. Model summary. A dysbiotic lung microbiota with lower bacterial diversity driven by colonization of potential respiratory pathogens such as *Haemophilus* and *Moraxella* precedes bacterial exacerbations in COPD patients. Elevated serum G-CSF at exacerbation classifies exacerbations characterized by bacteria (increased abundance of respiratory pathogens such as *Haemophilus* and *Moraxella,* increased bacterial load, and reduced lung microbiome diversity) and neutrophilic inflammation (high sputum and blood neutrophils and sputum IL-1β).