medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Circulating fibrocytes can identify people with a poor prognosis in pulmonary fibrosis

Iain D. Stewart[1,2], Henry Nanji[3], Grazziela Figueredo[3], William A. Fahy[4], Toby M. Maher[5,6], Antje J. Ask[7], Shyam Maharaj[7], Kjetil Ask[7], Martin Kolb[7]*, Gisli R. Jenkins[1,2]*.

1. Division of Respiratory Medicine, School of Medicine, University of Nottingham, UK
2. NIHR Biomedical Research Centre, University of Nottingham, UK
3. Advanced Data Analysis Centre, School of Computer Science, University of Nottingham, UK
4. Discovery Medicine, GlaxoSmithKline R&D, GlaxoSmithKline Medicines Research Centre, Stevenage, UK.
5. National Heart and Lung Institute, Imperial College, London, UK
6. National Institute of Health Research, Clinical Research Facility, Royal Brompton Hospital, London, UK.
7. Division of Respirology, Department of Medicine, McMaster University, Canada

*Authors contributed equally to this manuscript

Corresponding Author

Dr Iain Stewart, B19 Clinical Sciences Building, City Hospital, Nottingham, NG5 1PB

Contributions: IS contributed to design, data acquisition, analysis and preparation of the original text. HN and GF contributed to analysis and design. AM, SM, KA, MK contributed to data acquisition and critical appraisal of the text. WF, TM contributed to study design, data acquisition and critical appraisal of the text. GJ conceived the study, contributed to study design, data acquisition and preparation of the original text.

This article has an online data supplement

Abstract word count: 245

Main Text word count: 2293

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Abstract

Rationale: Circulating fibrocytes are elevated in idiopathic pulmonary fibrosis, but the relationship between fibrocyte level with lung function decline and clinical outcomes remains controversial.

Objective: To provide evidence for the utility of circulating fibrocyte levels as a prognostic biomarker in pulmonary fibrosis.

Methods: We tested associations between circulating fibrocyte levels, mortality, disease progression and longitudinal lung function in a well-defined prospective observational study of pulmonary fibrosis (PROFILE; NCT01134822). A subset of recruited participants had blood samples processed for fibrocyte measurement, with flow cytometry based on CD45 and collagen-I gating. Associations were tested using univariate and multivariate generalised linear models. Mortality data were combined with an independent study in a weighted-pooled analysis.

Measurements and Main Results: In 119 participants with idiopathic pulmonary fibrosis and non-specific interstitial pneumonia, circulating fibrocytes were elevated but not significantly above healthy controls. In multivariate analyses, each increment in circulating fibrocyte was associated with a 4% increased risk of overall mortality (Hazard Ratio 1.04 95% CI 1.00-1.09). A weighted-pooled analysis of idiopathic pulmonary fibrosis alone supported these observations in a larger sample of 162 participants, a 5% rise in fibrocytes was associated with a 24% increase in adjusted mortality risk (Hazard Ratio 1.24 95% CI 1.06-1.46). We found no association of fibrocytes with lung function or disease progression.

Conclusions: In a large prospective clinical cohort of pulmonary fibrosis, circulating fibrocytes can be isolated and increased levels are associated with increased mortality but do not associate with disease related decline in lung function.

Keywords: pulmonary fibrosis, biomarkers, mortality

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Introduction

Pulmonary fibrosis is the process of progressive scarring of lung tissue due to dysregulation of extracellular matrix turnover at the site of wound healing, the hallmark feature of a number of respiratory diseases.[1] Idiopathic pulmonary fibrosis (IPF) is a particularly severe and well-characterised type of pulmonary fibrosis with a 3-year median survival and an increasing incidence globally.[2] Individuals with IPF have varying degrees of fibrosis at presentation, and heterogeneneous progression of their illness. Understanding the mechanisms underpinning this variability may provide new therapeutic avenues or biomarkers of disease.

Fibrocytes are circulating mesenchymal progenitor cells that differentiate into tissue specific fibroblasts and contribute to multiple wound healing processes, including secretion of inflammatory cytokines, contractile wound closure, and promotion of angiogenesis.[3] The mesenchymal properties of the fibrocyte enables tissue invasion where it is thought they have pathogenic roles in fibrosis through transdifferentiation and accumulation of activated fibroblasts without resolution, or via paracrine signalling to resident cell types ultimately leading to production of dysregulated collagens through TGF-β mediated pathways.[4]

The role of fibrocytes in IPF remains controversial. Studies in transgenic mouse models have demonstrated that fibrocytes are not a necessary source of collagen-I and do not contribute to aberrant deposition in lung fibrosis.[5] However, a number of putative links between fibrocytes and disease activity in IPF have been reported, including a clinical study observing higher proportions during IPF exacerbation.[6] Thus, the contribution of fibrocytes to the pathogenesis of progressive pulmonary fibrosis remains unclear.

To improve our understanding of the role of fibrocytes in the pathogenesis of progressive pulmonary fibrosis we utilised the PROFILE study, a prospective multi-centre longitudinal cohort of patients with progressive fibrotic Interstitial Lung Diseases including IPF and fibrotic NSIP,[7 8] to investigated whether circulating fibrocytes can be reliably measured and whether they associated with key clinical outcomes including overall mortality, disease progression and lung function decline.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Methods

Participants and Study Design

Individuals with IPF or idiopathic NSIP confirmed by multidisciplinary team assessment were recruited to the PROFILE study as previously described.[7 8] Blood samples were collected at baseline and six months and processed for circulating fibrocyte numbers in 119 PROFILE participants who were representative of the entire cohort (Supplementary Table 1); samples from age-matched healthy controls were also collected for comparison. Individuals in the PROFILE study had demographic data and lung function recorded with twelve months of clinical follow up available. Disease progression was defined as a 10% relative decline in FVC or death within 12 months, survival was censored at 31 March 2016. Data were subsequently pooled with independent circulating fibrocyte data from individual IPF patients (Moeller et al. 2009),[6] diagnosed using the same standards.[9]

Fibrocyte Processing and Measurement

Fresh blood samples were centrifuged in Accuspin System Histopaque-1077 tubes at 1000x g for 10 mins. The mononuclear cell layer was transferred and resuspended before storage in liquid nitrogen, and shipped as a single batch on dry-ice. Fibrocytes were measured as described in Moeller et al.[6] Briefly, cells were first immunostained for the surface antigen CD45 (CD45-PerCP, #557513; BD Biosciences) or isotype control antibody (IgG_1, #555751; BD Biosciences), and permeabilized for detection of intracellular antigens. Cells were incubated with specific rabbit anti–human collagen-1 antibody (#6004011030.1 Rockland Immunochemicals) or IgG isotype control antibody (#AB105C; R&D Systems) followed with secondary antibody conjugated to Alexa Fluor 488 (#A20981; Invitrogen). Flow cytometry was performed with FACSDiva software (BD Biosciences). All data were analyzed with FlowJo software (Tree Star, Inc., Ashland, OR). The negative threshold for CD45 was set at 0.5%, using unstained cells, and all subsequent samples were gated for the CD45+ region. Cells gated for CD45 were

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

analyzed for collagen-1 expression, with negative control thresholds set at 0.5%, using a matched IgG isotype control. Specific staining for collagen-1 was determined as an increase in positive events over this threshold. Fibrocyte levels are defined as the percentage of circulating fibrocytes in total circulating leukocytes.

Statistical Analysis

Changes in paired fibrocyte levels at baseline and six months were compared where available (Supplementary Table 2). There were no measurable or statistically significant differences between paired values, and this provided justification to use the first chronologically occurring fibrocyte measurement as baseline for all individuals to maximise sample size. Fibrocyte levels were compared between IPF, NSIP and healthy controls using nonparametric tests, addressing disease status.

Cox proportional hazard models assessed univariate and multivariate adjusted overall mortality risk, sensitivity analyses were performed censoring at 12 months or with additional adjustment for background therapy. Proportional hazard results are presented with 95% confidence intervals adjusted for robust standard errors, and with adjustment for *a priori* confounders of baseline FVC, age, and gender. Kaplan-Meier survival analyses were performed at a threshold of 5%,[6] assessed by log-rank test. Kaplan-Meier estimates of potential follow-up (KM-PF) are used to describe distribution of follow-up time. Cox proportional hazard models in the pooled analysis for 5% increments in fibrocytes were fit using fixed effect inverse variance weighting. Heterogeneity between studies was measured by Cochran's Q, I^2, DerSimonian-Laird estimates of tau^2 and modified H^2.

Associations with lung function metrics FVC, DLCO and FEV1/FVC ratio at baseline and change over time were assessed according to fibrocyte level in linear models. To account for any potential hierarchical data structure, clusters were selected based on highest Calinski/Harabasz index, determined by k-means algorithm using baseline FVC percent predicted, fibrocyte level, age, smoking status and gender as classifiers, and analysed with mixed effects regression models with adjustment for clusters. Differences in fibrocyte level according to disease progression status was assessed by t-test. Statistical

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

significance was defined as $p<0.05$ throughout. All analyses were performed with Stata 16.0 (StataCorp TM).

Results

In total, 102 (85.7%) participants sampled for fibrocyte measurement were diagnosed with IPF and 17 (14.3%) participants were diagnosed with NSIP. Baseline characteristics of NSIP and IPF patients were similar (Table 1). On average, the total cohort with pulmonary fibrosis was 73 years of age, 75% male and 75% ever smokers. FVC was 82.7% of predicted value on average, whilst DLCO was 49.3% of predicted value; FEV1/FVC was 80%.

The percentage of circulating fibrocytes was highly variable across all participant groups. Median circulating fibrocyte level was higher in NSIP (3.84%) than healthy controls (1.61%), whilst levels were highest in IPF (4.49%), but these were not significantly different (Figure 1). Median levels in the entire disease group (4.37%) were not statistically significantly higher than in healthy controls ($p=0.073$). The range of fibrocyte levels extended furthest in IPF (38.6%), compared with NSIP (24.1%) and controls (16.6%).

In survival analyses, participants were followed up for a median 38.1 months (IQR 25.8–44.1). Univariate analysis of fibrocyte level in participants with IPF indicated that each percent increment in fibrocyte level was associated with a 3% increase in risk of overall mortality (HR 1.03 95%CI 1.00-1.07) and a 4% increase in risk where IPF and NSIP were combined, which remained significant following adjustment (Table 2). Overall, median fibrocyte level in participants who survived follow up was 3.7% (IQR 1.2%-8.1%), whilst for those who died (37%) it was 5.4% (IQR 2.4%-11.2%) ($p=0.074$).

Previous study has identified a 5% threshold of circulating fibrocytes as significantly associated with greater risk of mortality,[6] which we did not observe in the present data alone ($p=0.165$) but was observed when study datasets were pooled in an unweighted survival analysis ($p=0.029$) (Figure 2). To maximise power in evaluating the prognostic

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

capacity of circulating fibrocytes, we performed weighted-pooled analysis of PROFILE data with data from Moeller *et al.* (2009) for people with IPF specifically (usual interstitial pneumonia pattern). Scaled to a 5% increment in circulating fibrocytes, each increase was associated with a 29% (HR 1.29 95%CI 1.13-1.47) and 24% (HR 1.24 95%CI 1.06-1.46) increased risk of death in univariate and multivariate analyses, respectively (Table 3). Risk was similar in a sensitivity analysis performed for 12 month mortality (HR 1.35 95%CI 1.02-1.78) (Supplementary Table 3), or additionally adjusted for steroid and immunosuppression therapy (HR 1.22 95%CI 1.02-1.46) (Supplementary Table 4).

No associations between fibrocyte level and lung function metrics were observed in linear models or with mixed effects adjusted for hierarchical data clustering (Table 4). Adjustment for *a priori* confounders of age, gender and smoking did not alter these findings. Similarly, fibrocyte levels were not found to be a significant predictor of lung function within clusters. The effect sizes of fibrocyte level on 12 month change in FVC and DLCO values were very small and non-significant (data not shown).

Median fibrocyte levels were not significantly elevated in participants with pulmonary fibrosis whose disease progressed within 12 months (4.9% IQR 2.0%-10.1%), compared with those having stable disease (3.7% IQR 1.1%-8.1%) (p=0.281) (Figure 3). Similar results were observed in a sensitivity analysis of IPF diagnosed participants alone (p=0.447).

Discussion

Using prospectively sampled patients recruited into a well-characterised, multicentre, clinical, longitudinal, observational study, we provide data that support the findings of smaller studies but add caution to the clinical utility of circulating fibrocyte levels. We demonstrate that whilst circulating fibrocyte levels are elevated in fibrotic lung disease compared with healthy controls, variability across participant groups limits diagnostic potential. Prognostic value was confirmed as increments in the proportion of circulating

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

fibrocytes were associated with greater risk of mortality in adjusted models of pulmonary fibrosis broadly, as well as in sensitivity and pooled analyses of IPF alone. Variability across participant groups, and a lack of an observable association with lung function metrics or disease progression, indicates that circulating fibrocytes offer sensitivity for poor clinical outcomes but poor specificity in distinguishing severity of pulmonary fibrosis.

We observe that each increment in circulating fibrocyte proportion was associated with a 4% increased risk of overall mortality, independent of age, gender, smoking status or lung function. Moeller and colleagues identified a significant drop in median survival where fibrocyte levels were greater than 5%.[6] Using the published threshold, we did not confirm a significant drop in median survival in our dataset. This discrepancy may be explained by the inclusion of samples from acute exacerbations of IPF within the independent dataset, which limits wider applicability of findings from the small cohort.[6] Combining the datasets together to enhance power in a weighted-pooled analysis for people diagnosed with IPF identified a 24% increased risk in all-cause mortality for fibrocyte increments of 5%, following adjustment for confounders. Overall, these findings support the hypothesis that fibrocytes are associated with a poor prognosis and may play a pathogenic role in progressive pulmonary fibrosis.

Elevated levels of circulating fibrocytes have been observed in IPF compared with healthy controls.[6 10-12] We observe variable levels in all groups including those considered healthy, whilst no difference in fibrocyte levels were observed according to disease progression. We find no evidence of circulating fibrocyte dynamics over six months, or a relationship to lung function at baseline or longitudinally, including within disease clusters. The lack of an observable relationship with lung function or disease activity directly contrasts with published findings from a small sample of IPF patients.[11] Correlations previously observed in RA-ILD are similarly limited by small numbers of samples.[13] These data suggest that circulating fibrocytes do not provide a specific measure of pulmonary fibrosis disease activity, and may be elevated in comorbidities not captured within the dataset, such as hypertension.[14]

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

We followed established methods to process and measure fibrocytes, however gating strategies based on CD45/collagen-I have been found to include some contamination of granulocytes,[10] and the differentiation stage of circulating fibrocytes from precursor monocytes may reflect discrete functional subsets.[15-17] Ongoing efforts to characterise lung-specific fibrocytes may overcome limitations of contamination and offer further insight into the prognostic potential of specific subtypes of fibroblast precursor cells.[10 18] Periostin is involved in differentiation of lung specific fibroblasts and associations with disease progression have been observed in lung fibrosis.[19-21] Further studies to distinguish the precise cellular lineages of circulating fibrocytes are required.

Several limitations were present in these data. A total of 51 individuals had complete fibrocyte data at both baseline and six months, and limited molecular profiling assays for healthy participants prevented further distinction based on functional subgroups. We tested longitudinal differences in fibrocytes where complete, finding minimal change and subsequent rationale for analysis of first chronological instance of fibrocyte level. Minimal change over time in disease groups suggests circulating fibrocytes may not be a useful longitudinal metric of short-term disease severity, whilst missingness reflects logistical difficulties in processing and measuring fibrocytes from clinical samples. We tested associatons independently and in pooled analyses with historic data, a relatively large sample of IPF patients with fibrocyte data measured using a consistent strategy.[6] PROFILE participants were naïve to anti-fibrotics during the study, which also makes them comparable with the historic cohort in the context of disease management. Additional adjustments for steroid therapy and immunosuppression did not substantially change findings, but naivety to anti-fibrotics restricts evaluation of circulating fibrocytes as a biomarker for clinical response to such therapy. Recently nintedanib, a licensed treatment for IPF, has been demonstrated to attenuate fibrocyte migration and fibroblast differentiation in vitro and in preclinical animal models.[4 20 22] Studies determining whether nintedanib affects circulating fibrocytes levels would be better able to answer this question.

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

We performed analysis in a prospective observational sample, naïve to therapy and representative of the wider PROFILE cohort, assessing pulmonary fibrosis broadly and with sensitivity for IPF. We extend findings from previous studies and pool IPF data to provide greater power for evidence of an association with overall mortality. To our knowledge, these results include the largest prospective IPF sample studied and are applicable to real-world incidence of individuals with pulmonary fibrosis.

The value of circulating fibrocytes in prognosticating pulmonary fibrosis is limited due to the technical challenges measuring fibrocytes and lack of dynamic change or association with short term disease behaviour. However further characterisation of tissue-specific fibrocytes within prospective interventional clinical trials may offer more accurate assessment of their potential as biomarkers of disease activity.

References

1. Maharaj S, Shimbori C, Kolb M. Fibrocytes in pulmonary fibrosis: a brief synopsis. *Eur Respir Rev* 2013;22(130):552-7. doi: 10.1183/09059180.00007713
2. Hutchinson J, Fogarty A, Hubbard R, et al. Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur Respir J* 2015;46(3):795-806. doi: 10.1183/09031936.00185114
3. Keeley EC, Mehrad B, Strieter RM. Fibrocytes: bringing new insights into mechanisms of inflammation and fibrosis. *Int J Biochem Cell Biol* 2010;42(4):535-42. doi: 10.1016/j.biocel.2009.10.014
4. Wollin L, Wex E, Pautsch A, et al. Mode of action of nintedanib in the treatment of idiopathic pulmonary fibrosis. *Eur Respir J* 2015;45(5):1434-45. doi: 10.1183/09031936.00174914
5. Kleaveland KR, Velikoff M, Yang J, et al. Fibrocytes are not an essential source of type I collagen during lung fibrosis. *J Immunol* 2014;193(10):5229-39. doi: 10.4049/jimmunol.1400753
6. Moeller A, Gilpin SE, Ask K, et al. Circulating fibrocytes are an indicator of poor prognosis in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2009;179(7):588-94. doi: 10.1164/rccm.200810-1534OC
7. Jenkins RG, Simpson JK, Saini G, et al. Longitudinal change in collagen degradation biomarkers in idiopathic pulmonary fibrosis: an analysis from the prospective, multicentre PROFILE study. *The Lancet Respiratory Medicine* 2015;3(6):462-72. doi: 10.1016/s2213-2600(15)00048-x
8. Maher TM, Oballa E, Simpson JK, et al. An epithelial biomarker signature for idiopathic pulmonary fibrosis: an analysis from the multicentre PROFILE cohort study. *The*

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Lancet Respiratory Medicine 2017;5(12):946-55. doi: 10.1016/s2213-2600(17)30430-7

9. American Thoracic S, European Respiratory S. American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. . *Am J Respir Crit Care Med* 2002;165(2):277-304. doi: 10.1164/ajrccm.165.2.ats01
10. Heukels P, van Hulst JAC, van Nimwegen M, et al. Fibrocytes are increased in lung and peripheral blood of patients with idiopathic pulmonary fibrosis. *Respir Res* 2018;19(1):90. doi: 10.1186/s12931-018-0798-8
11. Fujiwara A, Kobayashi H, Masuya M, et al. Correlation between circulating fibrocytes, and activity and progression of interstitial lung diseases. *Respirology* 2012;17(4):693-8. doi: 10.1111/j.1440-1843.2012.02167.x
12. Mehrad B, Burdick MD, Zisman DA, et al. Circulating peripheral blood fibrocytes in human fibrotic interstitial lung disease. *Biochem Biophys Res Commun* 2007;353(1):104-8. doi: 10.1016/j.bbrc.2006.11.149
13. Just SA, Lindegaard H, Hejbol EK, et al. Fibrocyte measurement in peripheral blood correlates with number of cultured mature fibrocytes in vitro and is a potential biomarker for interstitial lung disease in Rheumatoid Arthritis. *Respir Res* 2017;18(1):141. doi: 10.1186/s12931-017-0623-9
14. Keeley EC, Mehrad B, Janardhanan R, et al. Elevated circulating fibrocyte levels in patients with hypertensive heart disease. *J Hypertens* 2012;30(9):1856-61. doi: 10.1097/HJH.0b013e32835639bb
15. Abe R, Donnelly SC, Peng T, et al. Peripheral blood fibrocytes: differentiation pathway and migration to wound sites. *J Immunol* 2001;166(12):7556-62. doi: 10.4049/jimmunol.166.12.7556
16. Wright AK, Brightling CE. Circulating fibrocytes: Will the real fibrocyte please stand up? *J Allergy Clin Immunol* 2016;137(5):1625-6. doi: 10.1016/j.jaci.2016.01.023
17. Pilling D, Fan T, Huang D, et al. Identification of markers that distinguish monocyte-derived fibrocytes from monocytes, macrophages, and fibroblasts. *PLoS One* 2009;4(10):e7475. doi: 10.1371/journal.pone.0007475
18. Reese C, Lee R, Bonner M, et al. Fibrocytes in the fibrotic lung: altered phenotype detected by flow cytometry. *Front Pharmacol* 2014;5:141. doi: 10.3389/fphar.2014.00141
19. Ashley SL, Wilke CA, Kim KK, et al. Periostin regulates fibrocyte function to promote myofibroblast differentiation and lung fibrosis. *Mucosal Immunol* 2017;10(2):341-51. doi: 10.1038/mi.2016.61
20. Sato S, Shinohara S, Hayashi S, et al. Anti-fibrotic efficacy of nintedanib in pulmonary fibrosis via the inhibition of fibrocyte activity. *Respir Res* 2017;18(1):172. doi: 10.1186/s12931-017-0654-2
21. Naik PK, Bozyk PD, Bentley JK, et al. Periostin promotes fibrosis and predicts progression in patients with idiopathic pulmonary fibrosis. *Am J Physiol Lung Cell Mol Physiol* 2012;303(12):L1046-56. doi: 10.1152/ajplung.00139.2012
22. Kasam RK, Reddy GB, Jegga AG, et al. Dysregulation of Mesenchymal Cell Survival Pathways in Severe Fibrotic Lung Disease: The Effect of Nintedanib Therapy. *Front Pharmacol* 2019;10:532. doi: 10.3389/fphar.2019.00532

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Figure Legends

Figure 1. Median circulating fibrocyte level by disease status. Fibrocyte levels (%) plotted according to disease diagnosis of NSIP or IPF, or from healthy age matched control. Median and interquartile range are presented for healthy (1.61 IQR: 0.63-4.66), NSIP (3.84 IQR: 1.64-10.20), and IPF (4.49 IQR 1.22-8.80). P-values provided for the difference between Healthy and NSIP (p=0.091), the difference between Healthy and IPF (p=0.086).

Figure 2. Kaplan Meier survival curve for threshold of circulating fibrocyte level in IPF. Proportion surviving plotted against follow up time in months. Threshold defined at 5% circulating fibrocytes; below threshold blue line, above threshold red line. A) Unweighted pooled PROFILE and Moeller *et al.* 2009 participants, log-rank p=0.029; B) PROFILE participants only, log-rank p=0.165.

Figure 3. Median circulating fibrocyte level in stable disease and progression. Fibrocyte levels (%) plotted according to disease progression (10% relative decline in FVC or mortality in 12 months), for NSIP and IPF together (A; p=0.281) and IPF alone (B; p=0.447).

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Table 1. Baseline characteristics according to disease status

	NSIP	IPF	IPF+NSIP
n	17	102	119
Median Age (IQR)	74 (66-76)	73 (68-79)	73 (68-79)
Male %	76.5	74.5	75.6
Ever Smoker %	64.7	77.5	75.6
FVC %pred (SD)	81.9 (15.5)	82.8 (20.7)	82.7 (20.0)
DLCO %pred (SD)	54.4 (10.5)	48.5 (16.7)	49.3 (16.1)
FEV1/FVC % (SD)	81.0 (7.5)	79.8 (8.4)	80.0 (8.25)

Non-normally distributed data presented as median and interquartile range (IQR), normally distributed data presented as mean and standard deviation (SD). NSIP – non specific interstitial pneumonia; IPF – idiopathic pulmonary fibrosis; FVC – forced vital capacity; DLCO – diffusion capacity; FEV1/FVC – forced expired volume in 1 second as a percentage of FVC; %pred – percentage of predicted value.

Table 2. Associations with risk of overall mortality.

	Univariate: IPF only			Univariate: IPF+NSIP			Multivariate: IPF+NSIP		
	HR	95%CI	p-value	HR	95%CI	p-value	HR*	95%CI	p-value
Fibrocyte	**1.03**	**(1.00-1.07)**	**0.024**	**1.04**	**(1.01-1.07)**	**0.004**	**1.04**	**(1.00-1.09)**	**0.04**
DLCO %pred	**0.93**	**(0.91-0.96)**	**0.001**	**0.93**	**(0.91-0.95)**	**<0.001**	**0.94**	**(0.91-0.96)**	**<0.001**
FVC %pred	**0.96**	**(0.95-0.98)**	**<0.001**	**0.96**	**(0.95-0.97)**	**<0.001**	0.98	(0.97-1.01)	0.22
Age	1.00	(0.96-1.05)	0.94	1.02	(0.97-1.06)	0.33	1.01	(0.97-1.05)	0.64
Male	1.11	(0.53-2.30)	0.77	1.03	(0.53-2.02)	0.91	1.09	(0.43-2.76)	0.84
Smoker	1.03	(0.46-2.30)	0.95	0.92	(0.46-1.85)	0.82	1.07	(0.44-2.62)	0.86

**Hazard ratio (HR) for 1% increments of circulating fibrocyte level adjusted for gender, age, smoking, and lung function. Significant associations in bold.*

Table 3. IPF risk of overall mortality for increments of 5% in circulating fibrocytes: weighted-pooled analysis.

Study	n	died	HR	95%CI	p-value	weight	HR*	95%CI*	p-value*	weight*
Moeller	60	16	1.52	(1.20-1.93)		30.2	1.49	(1.04-2.15)		19.7
PROFILE	102	38	1.20	(1.02-1.40)		69.8	1.19	(0.99-1.42)		80.3
Overall	162	54	1.29	(1.13-1.47)	<0.001	100.0	1.24	(1.06-1.46)	0.009	100.0

**Hazard ratio (HR) for 5% increments of circulating fibrocyte level adjusted for gender, age and percent predicted FVC in fixed-effects model. Heterogeneity in unadjusted model: Cochran's Q 2.64 p=0.105; I^2 62.1%; mH^2 1.635; tau^2 0.018. Heterogeneity in adjusted model: Cochran's Q 1.22 p=0.269; I^2 18.2%; mH^2 0.223; tau^2 0.005, n=158.*

Table 4. Mixed effect regression of cluster data for circulating fibrocyte level and lung function

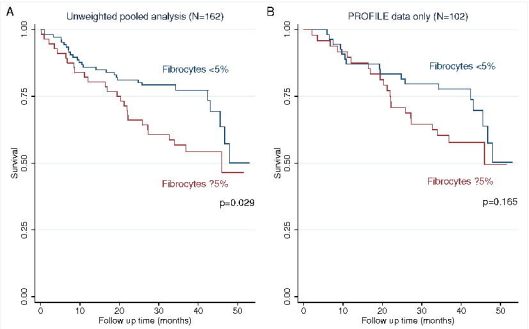
	Linear Modelling				
	n	Fibrocyte R^2	p-value	Fibrocyte* R^2	p-value
FVC %pred	118	0.8%	0.32	0.9%	0.45
DLCO %pred	104	0.6%	0.44	3.2%	0.47
FEV1/FVC	118	0.2%	0.62	3.3%	0.70
	Mixed effects for clustered data (clusters n=4)				
	n	Fibrocyte coefficient (95%CI)	Fibrocyte* coefficient (95%CI)	σ_u ; σ_e	ICC (%)
FVC %pred	110	3.03 (-1.06, 7.12)	2.50 (-1.43, 6.50)	216.7; 200.1	52
DLCO %pred	97	-0.17 (-0.91, 0.57)	-0.31 (-1.09, 0.47)	128.0; 132.0	49
FEV1/FVC	110	-0.37 (-0.65, 0.08)	-0.33 (-0.65, 0.08)	55.9; 23.6	70

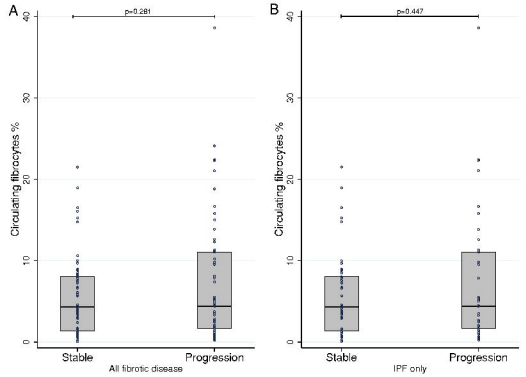
**adjusted for gender, age, smoking, and clusters where relevant. Four clusters were selected based on highest Calinski/Harabasz index, determed by k-means algorithm using baseline FVC %pred, fibrocyte level, age,*

medRxiv preprint doi: https://doi.org/10.1101/2020.06.05.20123406; this version posted June 7, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

smoking status and gender. σ_u between cluster variance; σ_e within cluster variance; ICC intraclass correlation coefficient.

p=0.066
p=0.091
Circulating Eosinophils %
0
10
20
30
40
Healthy
NSIP
IPF


A
Unweighted pooled analysis (N=162)
Fibrocytes <5%
Fibrocytes ≥5%
p=0.029
Survival
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
Follow up time (months)
B
PROFILE data only (N=102)
Fibrocytes <5%
Fibrocytes ≥5%
p=0.165
Survival
1.00
0.75
0.50
0.25
0.00
0
10
20
30
40
50
Follow up time (months)


A
p=0.281
Circulating fibrocytes %
0
10
20
30
40
Stable
Progression
All fibrotic disease
B
p=0.447
Circulating fibrocytes %
0
10
20
30
40
Stable
Progression
IPF only