

Longitudinal functional changes with clinically significant radiographic progression in idiopathic pulmonary fibrosis: are we following the right parameters?

Nada Taha¹, Dejanira D'Amato², Karishma Hosein¹, Tiziana Ranalli², Gianluigi Sergiacomi², Maurizio Zompatori³, Marco Mura¹

¹Division of Respiriology, Western University, London, Ontario, Canada

²Diagnostica per Immagini e Radiologia Interventistica, Policlinico Tor Vergata, University of Rome "Tor Vergata", Rome, Italy

³Radiologia, MultiMedica Group, I.R.C.C.S. San Giuseppe Hospital, Milan, Italy

Corresponding author:

Marco Mura, MD, PhD

London Health Science Centre, Victoria Hospital

800 Commissioners Road East Room E6-203

N6A 5W9 London, Ontario, Canada

Tel. +1-519-667-6744 Fax +1-519-685-8406

e-mail marco.mura@lhsc.on.ca

Word count: 2,514

ABSTRACT

Background. Progression of the disease in idiopathic pulmonary fibrosis (IPF) is difficult to predict, due to its variable and heterogenous course. The relationship between radiographic progression and functional decline in IPF is unclear. We sought to confirm that a simple HRCT fibrosis visual score is a reliable predictor of mortality in IPF, when longitudinally followed; and to ascertain which pulmonary functional variables best reflect clinically significant radiographic progression.

Methods: One-hundred-twenty-three consecutive patients with IPF from 2 centers were followed for an average of 3 years. Longitudinal changes of HRCT fibrosis scores, forced vital capacity (FVC), total lung capacity and diffusing lung capacity for carbon monoxide were considered. HRCTs were scored by 2 chest radiologists. The primary outcome was lung transplant (LTx)-free survival after the follow-up HRCT.

Results: During the follow-up period, 43 deaths and 11 LTx occurred. On average, the HRCT fibrosis score increased significantly, and a longitudinal increase $>7\%$ predicted LTx-free survival significantly, with good specificity, but limited sensitivity. The correlation between radiographic and functional progression was significant, but only moderate. HRCT progression and FVC decline predicted LTx-free survival independently and significantly, with better sensitivity, but worse specificity for a $\geq 5\%$ decline of FVC. However, the area under the curve towards LTx-survival were only 0.61 and 0.62, respectively.

Conclusions: The HRCT fibrosis visual score is a reliable and responsive tool to detect clinically meaningful disease progression. Although no individual pulmonary function test closely reflects radiographic progression, a longitudinal FVC decline improves sensitivity in the detection of

clinically significant disease progression. However, the accuracy of these methods remains limited, and better prognostication models need to be found.

Key words: idiopathic pulmonary fibrosis, high resolution chest CT scan, lung function, fibrosis score, survival.

BACKGROUND

One of the major challenges in the management of idiopathic pulmonary fibrosis (IPF) is predicting the clinical course of the disease in an individual patient, given its variability and heterogeneity (1). The detection of clinically significant progression is paramount to take decisions in terms of prognostication, anti-fibrotic therapy, lung transplant (LTx) assessment and listing. The use of clinically meaningful endpoints is also decisive in the conduction of new clinical trials in IPF (2).

Progression of disease IPF usually occurs with gradual worsening of radiographic fibrotic changes and physiological studies, including pulmonary function tests (PFTs) (3) and 6-minute walk test (4), in addition to clinical symptoms, including dyspnea and cough. As a result, commonly used predictors of disease progression include forced vital capacity (FVC) decline (5) and the repeat use of high-resolution chest CT scan (HRCT)(6). The decline in the FVC has been in fact the driving parameter for the approval of anti-fibrotic drugs in IPF (7,8). This has been preferred over the use of HRCT as an endpoint, given the concerns over exposure to radiation and cost.

Several methods exist to evaluate the extent of lung fibrosis on HRCT and to assess disease progression. These include visual scores (6,9) and automated methods (10-13). One

recognized limit in the use of HRCT in interstitial lung disease is interobserver variability in terms of pattern recognition (14,15). Despite the fact that automated methods of fibrosis scoring on HRCT may not be readily available in the majority of institutions, only a surprisingly small number of studies investigated the use of HRCT fibrosis visual scores in the longitudinal assessment of IPF (16,17). At the same time, concerns have arisen on the reliability of FVC alone in determining therapeutic decisions. These include missing progression in patients with stable FVC (18), observed intra-subject variability (19) and the fact that FVC decline itself is slowed down by anti-fibrotic therapy (2).

Previous studies investigated the relationship between concomitant radiographic and functional progression in IPF, but they were limited by either short follow-up period (6,20-22), and/or by the inclusion of a small cohort (21-23). Moreover, none of the previous studies applied c-statistics to investigate the accuracy of radiographic vs. functional progression to predict outcome. Collinearity between functional and radiographic progression was not studied either.

We previously demonstrated that baseline HRCT fibrosis visual score predicted 3-year survival in a cohort of patients newly diagnosed with IPF significantly (24). It remains to be established whether longitudinal changes of an HRCT visual fibrosis score represent a reliable predictor of progression in IPF. Furthermore, the relationship between radiographic and functional progression in IPF remains to be ascertained.

The objectives of this multicenter study were therefore: 1) to confirm that a simple HRCT fibrosis visual score is a reliable predictor of mortality in IPF, when longitudinally followed; 2) to ascertain which pulmonary functional variables best reflect clinically significant progression of disease seen on HRCT; 3) to assess the accuracy of mortality prediction of both radiographic and functional progression.

METHODS

Subjects

Consecutive patients diagnosed with IPF were enrolled in local databases and followed by 2 centres, London, ON, Canada and Rome, Italy. The diagnosis of IPF was based on the ATS/ERS/JRS/ALAT guidelines (25) in the context of multi-disciplinary discussion (MDD) (26). Longitudinal data were retrospectively collected. Patients who had 2 consecutive HRCT and corresponding PFT within 3 months from the HRCT scan date, and at least 1 year of clinical follow up from the time of diagnosis were included in the study. Patients with any other type of interstitial lung disease were excluded.

The follow-up HRCT was ordered by the attending physician as clinically indicated. Variables considered at baseline and follow-up assessment included an HRCT fibrosis visual score and measures of lung function, including forced vital capacity (FVC), total lung capacity (TLC), and diffusing lung capacity for carbon monoxide (DLCO). Twenty-nine patients were not able to perform a DLCO either at baseline or at follow-up assessment. Nineteen patients were not able to perform a TLC either at baseline or at follow-up assessment. There were no other missing data. Pulmonary function tests were performed according to ERS/ATS guidelines (27-29).

The primary outcome of the study was lung transplant-free survival from the date of the follow-up HRCT. The study was approved by the research ethics boards of Western University (protocol n. 101386) and University of Rome “Tor Vergata” (protocol n.175/15).

HRCT

The overall extent of pulmonary fibrosis was evaluated with a quantitative visual score, as previously described (9,24). Two chest radiologists from different institutions, with specific interstitial lung disease expertise and >20 years experience (G.S., M.Z.), who were blinded to clinical information, separately assessed the pattern and scored each case. Interobserver variability was assessed, and the average score was calculated. HRCT images were assessed for the presence and extent of parenchymal abnormalities, including reticular opacities, ground-glass opacities, traction bronchiectasis and honeycombing. A 5-point scale (0 = absence of lesions, 1, 2, 3, 4 = extent of lesions, respectively, <25, 25-50, 50-75, >75%) was used to determine the extent of overall fibrotic lung disease. The scores assigned for each scan at 4 predefined levels (aortic arch, bronchus intermedius, pulmonary veins, lowest scan) and each hemithorax were summed, and a final value was obtained: score = 100/maximum predicted value (equivalent to 8 times the number of scans performed). The extent of fibrosis was expressed as a percentage of the total lung volume. Baseline and follow-up HRCT fibrosis visual scores were calculated.

Statistical analysis

Values are expressed as mean \pm standard deviation. HRCT fibrosis score intervals were determined as: 0-10% -1, 11-20% -2, 21-30% - 3, and here on. Weighted kappa coefficient was then used to calculate interobserver agreement in quantitative fibrosis scoring (30). The Kolmogorov-Smirnov test was used for distribution analysis. Spearman coefficients were then used to assess linear association between variables. We calculated the variance inflation factors of the predicting variables to rule out the possibility of multicollinearity and demonstrate that the variables were truly independent (31). Receiver operating characteristic (ROC) analysis (c-statistics) was used to determine best cut points of each variable towards the endpoints, by

examining accuracy (sum of sensitivity and specificity) of predicting endpoints. Univariate and multivariate Cox proportional hazards regression analyses were performed identifying the significance of variables predicting the endpoint. p values <0.05 were regarded as significant. GraphPad (MacKiev, San Diego, CA), JMP (SAS Institute, Cary, NC), MedCalc (MedCalc, Mariakerke, Belgium) and Stata (Stata Corp, College Station, TX) softwares were used.

RESULTS

One hundred twenty-three patients from the 2 centers were included in the study. Demographic, clinical and functional characteristics are shown in Table 1. Eighty-six patients (70%) were diagnosed with based on clinical-radiographic criteria, and thirty-seven (30%) based on clinical-radiographic-pathologic criteria. The average interval between the baseline and follow-up HRCTs was 17 ± 7 months and the average length of follow-up after the 2nd HRCT was 19 ± 9 months. The average total follow-up after diagnosis was 36 ± 14 months. Eighty-four percent of patients received anti-fibrotic therapy, for an average length of 24 ± 23 months.

Interobserver agreement between chest radiologists on HRCT fibrosing scoring was good, with a weighted kappa coefficient of 0.64 (S.E. 0.05, C.I. 0.54-0.75).

At the end of the follow-up period, 43 patients (35%) had died, 11 patients (9%) received a LTx and 69 patients (56%) were alive without a LTx. Therefore, the combined incidence of death/LTx was 44%.

Longitudinal radiographic and functional changes were considered (Figure 1). During the interval between baseline and follow-up HRCT, a significant increase of the HRCT fibrosis score was observed ($p<0.0001$). During the same interval, only a non-significant decline of FVC and TLC was observed, while the decline of DLCO was significant ($p=0.0366$).

The relationships between HRCT changes and lung function changes are shown in Figure 2. Although these were significant, they were rather weak, ranging between 0.20 and 0.35 of correlation factor. VIF analysis demonstrated low collinearity between radiographic progression and functional progression: VIF was 1.92 for HRCT fibrosis score vs FVC, 1.74 vs. TLC, and 1.44 vs. DLCO.

We then considered the predictors of survival after the follow-up HRCT (Table 2). Age, gender and BMI were not significant. HRCT fibrosis score at baseline was a significant risk factor towards mortality ($p=0.0057$). ROC analysis (Table 3) determined that a $>7\%$ increase of HRCT fibrosis score from baseline was the most accurate interval change to predict mortality. This indeed was significant on univariate regression analysis ($p=0.0070$, Table 2).

Baseline FVC, TLC and DLCO were significantly protective towards mortality at baseline, but only a $\geq 5\%$ longitudinal decline of FVC ($p=0.0091$) or DLCO ($p<0.0001$) predicted mortality significantly.

In terms of accuracy of mortality prediction, ROC analysis (Table 4) demonstrated that a $>7\%$ longitudinal increase of HRCT fibrosis score was the most specific parameter (71%), while a $\geq 5\%$ decline of DLCO was the most sensitive (70%). However, only HRCT fibrosis score $>7\%$ and FVC decline $\geq 5\%$ were actually significant. Areas under the curve for HRCT fibrosis score, FVC, TLC and DLCO longitudinal changes ranged between 0.56 and 0.62 (Table 4).

Multivariate regression analysis (Table 4) demonstrated that a $>7\%$ increase of HRCT fibrosis score and a $\geq 5\%$ decline of FVC were independently and significantly predictive of mortality ($p=0.0260$ and $p=0.0152$, respectively).

DISCUSSION

In this study, we first demonstrated that a simple and reproducible HRCT visual fibrosis score is reproducible among chest radiologists with ILD expertise, and also a reliable and responsive predictor of mortality, when longitudinally assessed in IPF. We then demonstrated that no individual pulmonary test closely reflects clinically significant radiographic progression. Although our data showed that a longitudinal decline of FVC predicts worse survival independently from HRCT longitudinal assessment, they also demonstrated that the prognostic accuracy of these methods is far from being optimal.

With prolonged follow-up (average of 3 years) of patients from 2 different centers, we were able to capture a significant amount of progression of disease, as reflected by the high incidence (44%) of death/LTx following the 2nd HRCT, even in the era of systematic anti-fibrotic therapy. This ensured that the study was adequately powered in terms of endpoint events. Moreover, the population studied was not derived from a clinical trial, where trial-ineligible patients would not be included, limiting the representation of the general population of patients with IPF.

Using a variety of statistical approaches, we ascertained that a >7% increase of HRCT fibrosis score (% of lung volume) is moderately specific in predicting LTx-free survival in IPF. The very significant, longitudinal increase of the score, observed in the context of high incidence of deaths/LTx, reflects that the score is responsive of progression of disease. Finally, an increase of HRCT fibrosis score was strongly associated with mortality/LTx, which makes it highly clinically significant.

HRCT assessment remains the only validated biomarker in the diagnosis and follow-up management of patients with IPF (32). Hwang et al. previously demonstrated that progression of honeycombing is specifically associated with worse survival, not surprisingly (22). However,

most studies were limited by short follow-up (11) or were not powered enough, due to a limited number of deaths observed (10).

Walsh et al. extensively analyzed inter-observer agreement in terms of interstitial pattern recognition on HRCT (15). There are however no previous reports about agreement among radiologists in terms of HRCT quantitative visual fibrosis scoring. With a weight kappa coefficient of 0.64, agreement in this study was reassuringly good. However, such results would need to be confirmed in non-academic centers and among radiologists with less experience. At this point, the gold standard for HRCT pulmonary fibrosis quantitative scoring remains to be established.

The availability of a full set of longitudinal pulmonary function data obtained in concomitance to HRCT allowed us to compare radiographic and functional progression. Only moderate correlation was observed between radiographic progression and functional deterioration. Although this may seem concerning in terms of longitudinal functional assessment, it may also point to complementary role of pulmonary function tests, when coupled with radiographic assessment. Importantly, VIF calculation demonstrated very low collinearity (<2 for all PFTs)(31) between radiographic progression and functional decline.

Sverzellati et al. found that CALIPER-based HRCT changes and FVC decline are independently predictive of survival in IPF (17). We expanded these findings by assessing the accuracy of survival prediction of both radiographic and functional progression. We demonstrated that, while radiographic progression is quite specific (71%) to predict mortality, functional progression is relatively more sensitive (65%). On one side, the results of the HRCT visual fibrosis score are importantly reassuring, as automated methods of fibrosis quantification may not be readily available in the majority of centers. On the other hand, the overall accuracy of

either method is far from being optimal, with areas under the curve of only 0.61-0.62. This fundamental finding points to a strong need for better refinement of prognostic methods in patients with IPF.

A usual assumption is that HRCT is more sensitive than functional methods to detect progression of disease in IPF (10). Our finding of superior sensitivity of the FVC decline, compared to the HRCT fibrosis score, in predicting LTx-free survival is relatively surprising. This result can certainly be interpreted a limitation of a visual score, where difficult to detect active disease may be missed, while its physiologic impact may be captured by pulmonary function tests. Another limitation of the visual score is represented by the challenging distinction of pre-existing emphysema from honeycombing. From this perspective, we expect that new automated models of fibrosis scoring, such as AMFM (10), CALIPER (22,32) and DTA (20), will improve the sensitivity of HRCT in detecting progression of disease. On the other hand, clinical progression may be also related to other factors, such as the development of associated pulmonary hypertension, malnutrition or muscle deconditioning (34).

The prognostic results of DLCO may seem disappointing but, with advanced disease, not all patients with IPF are able to perform it (35). Additional limitations of DLCO measurement include technical variability, need for adjustment in certain conditions such as anemia, and inconsistency when patients are tested in different laboratories (35).

The major limitation of this study is represented by the use of a visual HRCT fibrosis score, which has well known limitations in terms of interobserver agreement and detection of small fibrotic changes, reflected by the low sensitivity of the HRCT visual score on ROC analysis in this study (36). On the other hand, new texture mapping softwares may not

immediately available to all centers, nor a specific automated method of fibrosis scoring has been validated yet.

Conclusions

The HRCT visual fibrosis score is a simple and relatively reliable tool to predict LTx-free survival in IPF, when longitudinally assessed, and with good interobserver agreement. While no individual pulmonary function tests perfectly reflect clinically-significant progression of disease detected on HRCT, this study confirms that the assessment of longitudinal decline of FVC still provides integrated and independent information, enhancing the sensitivity of prognostic models, with no collinearity found. However, and concerning, the accuracy of both HRCT and FVC remains limited, and better prognostic models in IPF are needed. Future studies need to establish whether this can be achieved by using new automated CT scoring methods, or, alternatively, by using a multi-dimensional approach implementing new components, such as exercise capacity and dyspnea scores, with lung function.

LIST OF ABBREVIATIONS

CTD: connective tissue disease

DLCO: diffusion capacity of carbon monoxide

FVC: forced vital capacity

GERD: gastroesophageal reflux disease

GGO: ground glass opacity

HRCT: high resolution CT scan

ILD: interstitial lung disease

IPF: idiopathic pulmonary fibrosis

LTx: lung transplant

6MWT: 6-minute walking test

PFT: pulmonary function test

PH: pulmonary hypertension

TLC: total lung capacity

UIP: usual interstitial pneumonia

DECLARATIONS

- The study was approved by the Research Ethics Boards of Western University (protocol n. 101386) and University of Rome “Tor Vergata” (protocol n.175/15).
- No personal identifiers of patients are included in the study.
- The authors declare that they have no competing interests.
- This study was supported by the Canadian Pulmonary Fibrosis Foundation (recipient Dr. Marco Mura) and by the Western University Department of Medicine Research Award (recipient Dr. Marco Mura).
- The dataset analyzed during the current study is not publicly available due to privacy restrictions dictated by the Research Ethics Boards. Selected data may be available from the corresponding author on reasonable request.
- Authors' contributions: NT - study design, data collection, analysis and interpretation, manuscript writing; DDA - study design, data interpretation and collection; KH - data collection; TR - data collection GS: study design, data analysis and interpretation,

manuscript review; MZ - study design, data analysis and interpretation, manuscript review; MM - study design, data collection, analysis and interpretation, manuscript writing.

REFERENCES

- 1: Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis: *Am J Respir Crit Care Med.* 2011, 15;183(4):431-40.
- 2: Kaner RJ, Bajwa EK, El-Amine M, Gorina E, Gupta R, Lazarus HM, Luckhardt TR, Mouded M, Posada K, Richeldi L, Stauffer J, Tutuncu A, Martinez FJ: Design of idiopathic pulmonary fibrosis clinical trials in the era of approved therapies. *Am J Respir Crit Care Med.* 2019, 15;200(2):133-139.
- 3: Richeldi L, Ryerson CJ, Lee JS, Wolters PJ, Koth LL, Ley B, Elicker BM, Jones KD, King TE Jr, Ryu JH, Collard HR: Relative versus absolute change in forced vital capacity in idiopathic pulmonary fibrosis. *Thorax.* 2012, 67(5):407-11.
- 4: du Bois RM, Albera C, Bradford WZ, Costabel U, Leff JA, Noble PW, Sahn SA, Valeyre D, Weycker D, King TE Jr: 6-Minute walk distance is an independent predictor of mortality in patients with idiopathic pulmonary fibrosis. *Eur Respir J.* 2014, 43(5):1421-9.
- 5: du Bois RM, Nathan SD, Richeldi L, Schwarz MI, Noble PW: Idiopathic pulmonary fibrosis: lung function is a clinically meaningful endpoint for phase III trials. *Am J Respir Crit Care Med.* 2012, 15;186(8):712-5.
- 6: Lynch DA, Godwin JD, Safrin S, Starko KM, Hormel P, Brown KK, Raghu G, King TE Jr, Bradford WZ, Schwartz DA, Richard Webb W, Idiopathic Pulmonary Fibrosis Study Group:

High-resolution computed tomography in idiopathic pulmonary fibrosis, diagnosis and prognosis. *Am J Respir Crit Care Med.* 2005, 15;172(4):488-93.

7: King TE Jr, Bradford WZ, Castro-Bernardini S, Fagan EA, Glaspole I, Glassberg MK, Gorina E, Hopkins PM, Kardatzke D, Lancaster L, Lederer DJ, Nathan SD, Pereira CA, Sahn SA, Sussman R, Swigris JJ, Noble PW, ASCEND Study Group: A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med.* 2014, 29;370(22):2083-92.

8: Richeldi L, du Bois RM, Raghu G, Azuma A, Brown KK, Costabel U, Cottin V, Flaherty KR, Hansell DM, Inoue Y, Kim DS, Kolb M, Nicholson AG, Noble PW, Selman M, Taniguchi H, Brun M, Le Maulf F, Girard M, Stowasser S, Schlenker-Herceg R, Disse B, Collard HR, INPULSIS Trial Investigators: Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med.* 2014, 29;370(22):2071-82.

9: Yoo DK, Zompatori M, Barrile A, Rossi G, D'Amato D, Sergiacomi G, Rogliani P, Mura M: Associated pulmonary hypertension is an independent contributor to exercise intolerance in chronic fibrosing interstitial pneumonias. *Respiration.* 2018;96(6):543-551.

10: Salisbury ML, Lynch DA, van Beek EJ, Kazerooni EA, Guo J, Xia M, Murray S, Anstrom KJ, Yow E, Martinez FJ, Hoffman EA, Flaherty KR, IPFnet Investigators: Idiopathic pulmonary fibrosis: the association between the adaptive multiple features method and fibrosis outcomes. *Am J Respir Crit Care Med.* 2017, 1;195(7):921-929.

11: Humphries SM, Swigris JJ, Brown KK, Strand M, Gong Q, Sundry JS, Raghu G, Schwarz MI, Flaherty KR, Sood R, O'Riordan TG, Lynch DA: Quantitative high-resolution computed tomography fibrosis score, performance characteristics in idiopathic pulmonary fibrosis. *Eur Respir J.* 2018,17;52(3).

- 12: Jacob J, Bartholmai BJ, Rajagopalan S, Kokosi M, Nair A, Karwoski R, Raghunath SM, Walsh SL, Wells AU, Hansell DM: Automated quantitative computed tomography versus visual computed tomography scoring in idiopathic pulmonary fibrosis, validation against pulmonary function. *J Thorac Imaging*. 2016, 31(5):304-11.
- 13: Park HJ, Lee SM, Song JW, Lee SM, Oh SY, Kim N, Seo JB: Texture-based automated quantitative assessment of regional patterns on initial CT in patients with idiopathic pulmonary fibrosis, relationship to decline in forced vital capacity. *AJR Am J Roentgenol*. 2016, 207(5):976-983. Epub 2016 Aug 17. PubMed PMID: 27533069.
- 14: Watadani T, Sakai F, Johkoh T, Noma S, Akira M, Fujimoto K, Bankier AA, Lee KS, Müller NL, Song JW, Park JS, Lynch DA, Hansell DM, Remy-Jardin M, Franquet T, Sugiyama Y: Interobserver variability in the CT assessment of honeycombing in the lungs. *Radiology*. 2013, 266(3):936-44.
- 15: Walsh SL, Calandriello L, Sverzellati N, Wells AU, Hansell DM, UIP Observer Consort: Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT. *Thorax*. 2016, 71(1):45-51.
- 16: Flaherty KR, Mumford JA, Murray S, Kazerooni EA, Gross BH, Colby TV, Travis WD, Flint A, Toews GB, Lynch JP 3rd, Martinez FJ: Prognostic implications of physiologic and radiographic changes in idiopathic interstitial pneumonia. *Am J Respir Crit Care Med*. 2003, 1;168(5):543-8.
- 17: Sverzellati N, Silva M, Seletti V, Galeone C, Palmucci S, Piciucchi S, Vancheri C, Poletti V, Tomassetti S, Karwoski R, Bartholmai BJ: Stratification of long-term outcome in stable idiopathic pulmonary fibrosis by combining longitudinal computed tomography and forced vital capacity. *Eur Radiol*. 2020, Jan 31. doi: 10.1007/s00330-019-06619-5.

31.

18: King TE Jr, Safrin S, Starko KM, Brown KK, Noble PW, Raghu G, Schwartz DA: Analyses of efficacy end points in a controlled trial of interferon-gamma1b for idiopathic pulmonary fibrosis. *Chest*. 2005, 127(1):171-7.

19: Nathan SD, Albera C, Bradford WZ, Costabel U, du Bois RM, Fagan EA, Fishman RS, Glaspole I, Glassberg MK, Glasscock KF, King TE Jr, Lancaster L, Lederer DJ, Lin Z, Pereira CA, Swigris JJ, Valeyre D, Noble PW, Wells AU: Effect of continued treatment with pirfenidone following clinically meaningful declines in forced vital capacity, analysis of data from three phase 3 trials in patients with idiopathic pulmonary fibrosis. *Thorax*. 2016, 71(5):429-35.

20: Humphries SM, Yagihashi K, Huckleberry J, Rho BH, Schroeder JD, Strand M, Schwarz MI, Flaherty KR, Kazerooni EA, van Beek EJR, Lynch DA: Idiopathic pulmonary fibrosis: data-driven textural analysis of extent of fibrosis at baseline and 15-month follow-up. *Radiology*. 2017, 285(1):270-278.

21: Clukers J, Lanclus M, Mignot B, Van Holsbeke C, Roseman J, Porter S, Gorina E, Kouchakji E, Lipson KE, De Backer W, De Backer J: Quantitative CT analysis using functional imaging is superior in describing disease progression in idiopathic pulmonary fibrosis compared to forced vital capacity. *Respir Res*. 2018, 6;19(1):213.

22: Hwang JH, Misumi S, Curran-Everett D, Brown KK, Sahin H, Lynch DA: Longitudinal follow-up of fibrosing interstitial pneumonia: relationship between physiologic testing, computed tomography changes, and survival rate. *J Thorac Imaging*. 2011, 26(3):209-17.

23: Jacob J, Bartholmai BJ, Rajagopalan S, Kokosi M, Egashira R, Brun AL, Nair A, Walsh SLF, Karwoski R, Wells AU: Serial automated quantitative CT analysis in idiopathic pulmonary

fibrosis: functional correlations and comparison with changes in visual CT scores. *Eur Radiol.* 2018, 28(3):1318-1327.

24: Mura M, Porretta MA, Bargagli E, Sergiacomi G, Zompatori M, Sverzellati N, Taglieri A, Mezzasalma F, Rottoli P, Saltini C, Rogliani P: Predicting survival in newly diagnosed idiopathic pulmonary fibrosis: a 3-year prospective study. *Eur Respir J.* 2012, 40(1):101-9.

25: Raghu G, Remy-Jardin M, Myers JL, Richeldi L, Ryerson CJ, Lederer DJ, Behr J, Cottin V, Danoff SK, Morell F, Flaherty KR, Wells A, Martinez FJ, Azuma A, Bice TJ, Bouros D, Brown KK, Collard HR, Duggal A, Galvin L, Inoue Y, Jenkins RG, Johkoh T, Kazerooni EA, Kitaichi M, Knight SL, Mansour G, Nicholson AG, Pipavath SNJ, Buendía-Roldán I, Selman M, Travis WD, Walsh S, Wilson KC; American Thoracic Society, European Respiratory Society, Japanese Respiratory Society, and Latin American Thoracic Society: Diagnosis of idiopathic pulmonary fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *Am J Respir Crit Care Med.* 2018, 198(5):e44-e68.

26: Flaherty KR, King TE Jr, Raghu G, Lynch JP 3rd, Colby TV, Travis WD, Gross BH, Kazerooni EA, Toews GB, Long Q, Murray S, Lama VN, Gay SE, Martinez FJ: Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? *Am J Respir Crit Care Med.* 2004, 170(8):904-10.

27: Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Jensen R, Johnson DC, MacIntyre N, McKay R, Navajas D, Pedersen OF, Pellegrino R, Viegi G, Wanger J; ATS/ERS Task Force: Standardisation of spirometry. *Eur Respir J.* 2005, 26(2):319-38.

28: Wanger J, Clausen JL, Coates A, Pedersen OF, Brusasco V, Burgos F, Casaburi R, Crapo R, Enright P, van der Grinten CP, Gustafsson P, Hankinson J, Jensen R, Johnson D, Macintyre N,

McKay R, Miller MR, Navajas D, Pellegrino R, Viegi G: Standardisation of the measurement of lung volumes. *Eur Respir J*. 2005, 26(3):511-22.

29: Graham BL, Brusasco V, Burgos F, Cooper BG, Jensen R, Kendrick A, MacIntyre NR, Thompson BR, Wanger J: 2017 ERS/ATS standards for single-breath carbon monoxide uptake in the lung. *Eur Respir J*. 2017, 3;49(1).

30: Brennan P, Silman A: Statistical methods for assessing observer variability in clinical measures. *BMJ*. 1992, 6;304(6840):1491-4.

31: Slinker BY, Glantz SA: Multiple regression for physiological data analysis: the problem of multicollinearity. *Am J Phys*. 1985, 249:R1-R12.

32: Ley B. Counterpoint: should molecular and genetic biomarkers be used in the initial evaluation of patients with fibrotic ILD? No. *Chest*. 2019, 156(2):205-208.

33: Jacob J, Bartholmai BJ, Rajagopalan S, Kokosi M, Nair A, Karwoski R, Raghunath SM, Walsh SL, Wells AU, Hansell DM: Automated quantitative computed tomography versus visual computed tomography scoring in idiopathic pulmonary fibrosis: validation against pulmonary function. *J Thorac Imaging*. 2016, 31(5):304-11.

34: King CS, Nathan SD: Idiopathic pulmonary fibrosis: effects and optimal management of comorbidities. *Lancet Respir Med*. 2017, 5(1):72-84.

35: Pellegrino R, Viegi G, Brusasco V, Crapo RO, Burgos F, Casaburi R, Coates A, van der Grinten CP, Gustafsson P, Hankinson J, Jensen R, Johnson DC, MacIntyre N, McKay R, Miller MR, Navajas D, Pedersen OF, Wanger J: Interpretative strategies for lung function tests. *Eur Respir J*. 2005, 26(5):948-68.

36: Lee SM, Seo JB, Oh SY, Kim TH, Song JW, Lee SM, Kim N: Prediction of survival by texture-based automated quantitative assessment of regional disease patterns on CT in idiopathic pulmonary fibrosis. *Eur Radiol.* 2018, 28(3):1293-1300.

FIGURE LEGENDS

Figure 1. Longitudinal changes from baseline. A. HRCT visual fibrosis score. B. FVC % predicted. C. TLC % predicted. D. DLCO % predicted.

Figure 2. Relationship between HRCT visual fibrosis score longitudinal change and: A. FVC % pred longitudinal change; B: TLC % pred longitudinal change; C: DLCO % pred longitudinal change.

TABLES
TABLE 1. Demographic, clinical and radiographic characteristics of the patients at baseline.

Variable	Average \pm SD	Range
Age (years)	71 \pm 8	43-89
Gender (% male)	72	
BMI (kg/m²)	30 \pm 5	
Smoking (pack-years)	22 \pm 20	0 - 113
MDD diagnosis		
Clinical-radiographic-pathologic (%)	37	
Clinical-radiographic (%)	86	
FVC (% pred)	82 \pm 21	38 - 132
TLC (% pred)	73 \pm 16	40 - 118
DLCO (% pred)	50 \pm 18	17 - 97
Antifibrotic therapy (%)	84	
Pirfenidone (%)	39	
Nintedanib (%)	23	
Switched therapy (%)	22	
Duration of antifibrotic therapy if any (months)	24 \pm 23	1-82
Outcomes		
Alive without LTx (%)	69 (56)	
LTx (%)	11 (9)	
Dead (%)	43 (35)	

HRCT scan pattern		
Definite UIP (%)	49	
Probable UIP (%)	43	
Indeterminate for UIP (%)	5	
Alternative diagnosis (%)	26	
Concomitant emphysema (%)	46	

Abbreviations: BMI: body mass index, DLCO: diffusion capacity, FVC: forced vital capacity, HRCT: high resolution CT scan, LTx: lung transplant, MDD: multi-disciplinary diagnosis, TLC: total lung capacity, UIP: usual interstitial pneumonia,

Table 2. Univariate regression analysis: predictors of lung transplant-free survival after the 2nd HRCT.

Variable	Hazard ratio (C.I.)	p value
Age (years)	1.01 (0.98-1.05)	n.s.
Gender (male)	1.18 (0.65-2.24)	n.s.
BMI (kg/m ²)	0.96 (0.90-1.01)	n.s.
Baseline HRCT fibrosis score	1.03 (1.01-1.06)	0.0057
HRCT fibrosis score change >7%	2.10 (1.23-3.62)	0.0070
Baseline FVC	0.98 (0.97-0.99)	0.0126
FVC change ≥5%	2.13 (1.20-3.92)	0.0091
Baseline TLC	0.97 (0.95-0.99)	0.0081
TLC change ≥5%	0.99 (0.96-1.02)	n.s.
Baseline DLCO	0.94 (0.94-0.98)	<0.0001
DLCO change ≥5%	1.31 (0.66-2.78)	n.s.

Table 3. Receive operating characteristics analysis. Predictors of lung transplant-free survival after the 2nd HRCT.

Variable	Area under the curve	Sensitivity (C.I.)	Specificity (C.I.)	p value
HRCT fibrosis score change (7%) from baseline	0.61	52% (38-66)	71% (59-81)	0.0266
FVC change (5%) from baseline	0.62	65% (50-78)	60% (47-71)	0.0193
TLC change (5%) from baseline	0.56	55% (39-70)	56% (43-69)	n.s.
DLCO change (5%) from baseline	0.57	70% (53-84)	44% (31-58)	n.s.

Table 4. Multivariate regression analysis: independent, significant predictors of lung transplant-free survival after the 2nd HRCT.

Variable	Hazard ratio (C.I.)	p value
FVC change ≥5%	2.07 (1.15-3.87)	0.0152
HRCT fibrosis score change >7%	1.92 (1.09-3.41)	0.0260



