



# New radiological diagnostic criteria: impact on idiopathic pulmonary fibrosis diagnosis

*To the Editor:*

The American Thoracic Society (ATS), European Respiratory Society (ERS), Japanese Respiratory Society (JRS) and Latin American Thoracic Society (ALAT) recently released a new clinical practice guideline (ATS/ERS/JRS/ALAT2018) for idiopathic pulmonary fibrosis (IPF) with simultaneously proposed diagnostic criteria by the Fleischner Society [1, 2]. Both diagnostic algorithms agree on most diagnostic steps, with divergent recommendations on the position of surgical lung biopsy (SLB): ATS/ERS/JRS/ALAT2018 [2] recommends SLB in most patients with a probable usual interstitial pneumonia (UIP) pattern on high-resolution computed tomography (HRCT) (conditional recommendation), whereas the Fleischner Society proposes to forgo SLB in patients with a definite or probable UIP HRCT pattern, presenting in the right clinical context [3, 4]. We aimed to quantify the impact of the previous approach (ATS/ERS/JRS/ALAT2011) [5] and the two new diagnostic approaches [1, 2] on real-life clinical practice, with assessment of radiological interrater agreement, diagnostic test characteristics, and prognostic validity of the diverging radiological diagnoses for a multidisciplinary IPF diagnosis in our cohort.

We included patients with a multidisciplinary team (MDT) diagnosis of IPF from our cohort study [6], and a clinically relevant control group of patients with fibrosing interstitial lung diseases (chronic hypersensitivity pneumonitis (cHP) and unclassifiable interstitial lung disease (ILD)). Approval by the local ethics committee was obtained for data acquisition (Swiss Ethics Committee, Bern, KEK 246/15 PB\_2016-01524).

Three subspecialised thoracic radiologists reviewed the HRCT scans blinded to the initial classification, clinical diagnosis and patient characteristics. During the first read out, raters classified the patterns according to the 2011 criteria: 1) definitive UIP pattern; 2) possible UIP pattern; and 3) inconsistent with UIP pattern. 4 weeks later, the radiologists received a re-randomised case collection for a second interpretation round according to the Fleischner Society and subsequently to the ATS/ERS/JRS/ALAT2018 recommendations: 1) typical UIP pattern; 2) probable UIP pattern; 3) indeterminate for UIP; and 4) alternative diagnosis.

For this analysis, we compared the three radiological scenarios allowing an IPF diagnosis without SLB in most patients: a definite UIP pattern according to ATS/ERS/JRS/ALAT2011 and ATS/ERS/JRS/ALAT2018 recommendations, and a definite or probable UIP pattern according to Fleischner Society criteria. Interrater agreement between the three radiologists was assessed using Light-Kappa ( $\kappa$ ), which is an extension of Cohen's Kappa for more than two raters [7]. Every radiological pattern was dichotomised (e.g. definite UIP yes/no), and we considered a pattern to be present if 2–3 out of three radiologists scored the pattern with yes, otherwise the pattern was scored not to be present (no). Test characteristics with 95% confidence intervals of the radiological scenarios against the MDT diagnosis of IPF as a reference standard were determined, including sensitivity, specificity, positive and negative predictive values (PPV and NPV), and Youden's index which summarises the diagnostic test performance. We assessed the prognostic validity of the three radiological scenarios by estimation of their effect on time to death using Cox proportional hazards models.



@ERSpublications

**Divergent recommendations and guidelines exist for IPF. Comparison of existing guidelines indicates IPF was diagnosed with better sensitivity and preserved specificity with Fleischner recommendations compared with current clinical practice guidelines.** <http://bit.ly/2ozH9la>

**Cite this article as:** Funke-Chambour M, Guler SA, Geiser T, *et al.* New radiological diagnostic criteria: impact on idiopathic pulmonary fibrosis diagnosis. *Eur Respir J* 2019; 54: 1900905 [<https://doi.org/10.1183/13993003.00905-2019>].

We included 52 patients with IPF (MDT diagnosis) and 37 with non-IPF ILD. Baseline characteristics of patients with IPF and other ILDs were similar. Out of 89 patients in total, 70 (79%) were male, 56 (63%) were ever-smokers, the mean $\pm$ SD age was 68.4 $\pm$ 10.4 years; mean $\pm$ SD body mass index was 27.2 $\pm$ 4.8 kg·m<sup>-2</sup>, mean $\pm$ SD forced vital capacity was 68.4 $\pm$ 19.0% predicted and mean $\pm$ SD diffusing capacity of the lung for carbon monoxide ( $D_{LCO}$ ) was 50 $\pm$ 16.7% predicted. SLB was available in 19 (37%) IPF and 21 (57%) patients with other ILDs. During follow-up (mean 26.9 months (interquartile range 17.2–50.2 months)) 28 (54%) of all IPF and 2 (5.4%) of other ILD patients died.

The overall interrater agreement was good for ATS/ERS/JRS/ALAT2011 ( $\kappa$  0.61), and moderate for ATS/ERS/JRS/ALAT2018 and Fleischner ( $\kappa$  0.54 and 0.57, respectively).

Diagnostic test characteristics for the MDT-IPF diagnosis demonstrated a high specificity for all three radiological scenarios. Fleischner misclassified four out of 37 non-IPF cases as IPF, compared with three out of 37 and one out of 37 misclassifications with ATS/ERS/JRS/ALAT2018 and ATS/ERS/JRS/ALAT2011, respectively. Sensitivity was markedly higher for the Fleischner scenario with seven out of 52 missed IPF cases, compared with 20 out of 52 and 23 out of 52 missed IPF cases for ATS/ERS/JRS/ALAT2018 and ATS/ERS/JRS/ALAT2011, respectively. PPV and NPV were slightly higher for Fleischner than for ATS/ERS/JRS/ALAT2018, and of all the scenarios, Fleischner had the strongest overall test performance (table 1).

Patients with IPF had significantly worse survival: MDT-IPF diagnosis was associated with a 16-fold increased risk of death compared to cHP and unclassifiable ILD (95% CI 3.7–68.9). All radiological scenarios significantly correlated with survival in unadjusted analysis, with Fleischner scoring (definite and probable UIP) having the largest impact on mortality (HR 4.00, 95% CI 1.69–9.53, table 1). The Fleischner scenario remained the strongest correlate of mortality risk including with adjustment for potential confounders (age, sex, ever smoker and  $D_{LCO}$  % predicted (HR 3.76, 95% CI 1.22–11.6, model C-index 0.80)). Corresponding multivariate Cox regression models demonstrated the ATS/ERS/JRS/ALAT2011 scenario to be independently associated with survival (HR 2.87, 95% CI 1.06–7.78, model C-index 0.81), whereas ATS/ERS/JRS/ALAT2018 lost statistical significance in the adjusted model (HR 2.01, 95% CI 0.71–5.68, model C-index 0.80).

In the light of changing paradigms in IPF management, the clinical diagnostic approach to patients with suspected IPF has recently been discussed by the Fleischner Society and the ATS/ERS/JRS/ALAT, culminating in the publication of two, slightly different, new diagnostic strategies [1, 2]. We compared the clinical impact of the specific ATS/ERS/JRS/ALAT2011, ATS/ERS/JRS/ALAT2018, and Fleischner radiological scenarios that allow clinical IPF diagnosis without SLB in most patients. Fleischner differs from the ATS/ERS/JRS/ALAT algorithms by accepting not only definite, but also a probable UIP pattern for IPF diagnosis in the right clinical context. In our cohort of well-characterised IPF patients compared with a control group, we demonstrate a good diagnostic accuracy of the Fleischner approach for MDT diagnosis of IPF, with a prognostic discrimination that markedly strengthens its validity.

Comparable to our findings from ATS/ERS/JRS/ALAT2011 and 2018 scoring, previous studies reported definite UIP as a highly specific but less sensitive marker for diagnosis of IPF, with clinical trials including patients without a definite HRCT UIP pattern [8, 9]. In this cohort, including patients with a probable UIP pattern in the subgroup not requiring SLB in the diagnostic algorithm results in a significantly higher

TABLE 1 Diagnostic and prognostic performance of radiological scenarios

Radiological scenario	Test characteristics <sup>#</sup>				Survival (unadjusted)			
	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	J-index (95% CI)	HR (95% CI)	p-value	C-index
ATS/ERS/JRS/ALAT 2011 <sup>¶</sup>	0.56 (0.41–0.70)	0.97 (0.86–1.00)	0.97 (0.83–1.00)	0.61 (0.47–0.73)	0.53 (0.27–0.69)	3.21 (1.54–6.67)	0.001	0.66
ATS/ERS/JRS/ALAT 2018 <sup>¶</sup>	0.62 (0.47–0.75)	0.92 (0.78–0.98)	0.91 (0.77–0.98)	0.63 (0.49–0.76)	0.53 (0.25–0.73)	3.26 (1.55–6.85)	0.002	0.64
Fleischner <sup>*</sup>	0.87 (0.74–0.94)	0.89 (0.75–0.97)	0.92 (0.80–0.98)	0.82 (0.67–0.93)	0.75 (0.49–0.91)	4.00 (1.69–9.53)	0.002	0.66

<sup>#</sup>: diagnostic test characteristics with idiopathic pulmonary fibrosis multidisciplinary team diagnosis as the reference standard; <sup>¶</sup>: definite usual interstitial pneumonia (UIP) scored according to ATS/ERS/JRS/ALAT 2011 and ATS/ERS/JRS/ALAT 2018 guidelines; <sup>\*</sup>: definite and probable UIPs scored according to the recommendations of the Fleischner Society. C-index: Harrell's concordance statistic; HR: hazard ratio; J-index: Youden's J statistic; NPV: negative predictive value; PPV: positive predictive value.

sensitivity with only marginal loss in specificity (incremental misclassification of 2.7%). Applying the Fleischner algorithm, we further found an increase in PPV and NPV compared with ATS/ERS/JRS/ALAT2018, with the limitation that these characteristics are not generalisable to cohorts with different IPF prevalences [10]. Patients' course of disease and survival supports IPF diagnosis [11]. We demonstrate that although the MDT-IPF diagnosis remains the strongest predictor of survival, the three radiological scenarios significantly correlated with mortality. In contrast to ATS/ERS/JRS/ALAT2018, the Fleischner radiological scenario was associated with risk of death independent from clinical baseline characteristics. Beyond the prognostic importance, this strengthens the diagnostic validity of the combination of definite and probable HRCT UIP for clinical IPF diagnosis.

Our work confirms moderate interobserver agreement between expert radiologists for definite and probable UIP pattern according to previous and current guidelines [12, 13]. This subjective component in radiological diagnosis might be problematic if integrated in clinical decision making on further invasive diagnostic procedures. Semi-automated CT readings might address this issue in the future [14]. Regardless, the potential benefit of the added pathological information still needs to be balanced carefully against the risks of invasive procedures such as SLB, particularly in elderly patients with severely impaired pulmonary function or significant comorbidities. The Fleischner diagnostic approach might reduce the percentage of patients with suspected IPF needing an invasive procedure, and data from our cohort support its diagnostic and prognostic validity in clinical routine.

Future prospective studies are needed to validate different decision algorithms incorporating noninvasive and, if needed, invasive biomarkers with inclusion of patient preference in the process.

**Manuela Funke-Chambour<sup>1,2,4</sup>, Sabina A. Guler<sup>1,4</sup>, Thomas Geiser<sup>1,2</sup>, Andreas Christe<sup>3</sup>, Johannes Heverhagen<sup>3</sup>, Alexander Pöllinger<sup>3</sup>, Adrian Huber<sup>3</sup> and Lukas Ebner<sup>3</sup>**

<sup>1</sup>Dept of Pulmonary Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland. <sup>2</sup>Dept for BioMedical Research, University of Bern, Bern, Switzerland. <sup>3</sup>Dept of Diagnostic, Interventional and Pediatric Radiology, Inselspital, Bern University Hospital, University of Bern, Switzerland. <sup>4</sup>Both authors contributed equally.

Correspondence: Manuela Funke-Chambour. Dept of Pulmonary Medicine, Inselspital Bern, Bern 3010, Switzerland  
E-mail: manuela.funke-chambour@insel.ch

Received: 06 May 2019 | Accepted after revision: 28 June 2019

Conflict of interest: M. Funke-Chambour reports grants from Roche, Boehringer Ingelheim, and Intermune (limited grant for initial cohort study set up), during the conduct of the study. S.A. Guler has nothing to disclose. T. Geiser has nothing to disclose. A. Christe has nothing to disclose. J. Heverhagen has nothing to disclose. A. Pöllinger has nothing to disclose. A. Huber has nothing to disclose. L. Ebner has nothing to disclose.

Support statement: Boehringer Ingelheim, F. Hoffmann-La Roche and Intermune provided an unrestricted grant for the initial cohort study. Funding information for this article has been deposited with the Crossref Funder Registry.

## References

- Lynch DA, Sverzellati N, Travis WD, *et al.* Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. *Lancet Respir Med* 2018; 6: 138–153.
- Raghu G, Remy-Jardin M, Myers JL, *et al.* Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018; 198: e44–e68.
- Richeldi L, Wilson KC, Raghu G. Diagnosing idiopathic pulmonary fibrosis in 2018: bridging recommendations made by experts serving different societies. *Eur Respir J* 2018; 52: 1801485.
- Wells AU. IPF diagnosis: flexibility is a virtue. *Lancet Respir Med* 2018; 6: 735–737.
- Raghu G, Collard HR, Egan JJ, *et al.* ATS/ERS/JRS/ALAT Committee on Idiopathic Pulmonary Fibrosis. An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.
- Guler S, Zumstein P, Berezowska S, *et al.* Idiopathic pulmonary fibrosis in a Swiss interstitial lung disease reference centre. *Swiss Med Wkly* 2018; 148: w14577.
- Light RJ. Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychological Bulletin* 1971; 76: 365–377.
- Richeldi L, du Bois RM, Raghu G, *et al.* INPULSIS Trial Investigators. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2071–2082.
- Raghu G, Wells AU, Nicholson AG, *et al.* Effect of nintedanib in subgroups of idiopathic pulmonary fibrosis by diagnostic criteria. *Am J Respir Crit Care Med* 2017; 195: 78–85.
- Brownell R, Moua T, Henry TS, *et al.* The use of pretest probability increases the value of high-resolution CT in diagnosing usual interstitial pneumonia. *Thorax* 2017; 72: 424–429.
- Jo HE, Glaspole I, Goh N, *et al.* Implications of the diagnostic criteria of idiopathic pulmonary fibrosis in clinical practice: analysis from the Australian Idiopathic Pulmonary Fibrosis Registry. *Respirology* 2019; 24: 361–368.
- Walsh SL, Calandriello L, Sverzellati N, *et al.* Interobserver agreement for the ATS/ERS/JRS/LATS /JRS/ALAT criteria for a UIP pattern on CT. *Thorax* 2016; 71: 45–51.
- Walsh SL, Maher TM, Kolb M, *et al.* Diagnostic accuracy of a clinical diagnosis of idiopathic pulmonary fibrosis: an international case-cohort study. *Eur Respir J* 2017; 50: 1700936.

- 14 Anthimopoulos M, Christodoulidis S, Ebner L, *et al.* Semantic Segmentation of Pathological Lung Tissue with Dilated Fully Convolutional Networks. *IEEE J Biomed Health Inform* 2019; 23: 714–722.

Copyright ©ERS 2019