EUROPEAN RESPIRATORY *journal*

FLAGSHIP SCIENTIFIC JOURNAL OF ERS

Early View

Research letter

New radiological diagnostic criteria – impact on idiopathic pulmonary fibrosis diagnosis

Manuela Funke-Chambour, Sabina A. Guler, Thomas Geiser, Andreas Christe, Johannes Heverhagen, Alexander Pöllinger, Adrian Huber, Lukas Ebner

Please 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; in press (https://doi.org/10.1183/13993003.00905-2019).

To the Editor

New radiological diagnostic criteria – impact on idiopathic pulmonary fibrosis diagnosis

Manuela Funke-Chambour[*1,2], Sabina A. Guler*[*1], Thomas Geiser[1,2], Andreas Christe[3], Johannes Heverhagen[3], Alexander Pöllinger[3], Adrian Huber[3], Lukas Ebner[3]*

**Both authors contributed equally to this manuscript*

[1]*Department of Pulmonary Medicine, Inselspital, Bern University Hospital, University of Bern, Bern, Switzerland*
[2]*Department for BioMedical Research, University of Bern, Bern, Switzerland;*
[3]*Department of Diagnostic, Interventional and Pediatric Radiology, Inselspital, Bern University Hospital, University of Bern, Switzerland*

Introduction

The ATS/ERS/JRS/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 recommends SLB in most patients with 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 definite or probable UIP HRCT pattern, presenting in the right clinical context. [3,4] We aimed to quantify the impact of the previous (ATS/ERS/JRS/ALAT2011) [5] and the two new diagnostic approaches [1,2] on real-life clinical practice, with assessment of radiological inter-rater agreement, diagnostic test characteristics, and prognostic validity of the diverging radiological diagnoses for a multidisciplinary IPF diagnosis in our cohort.

Methods

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 subspecialized 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. Four weeks later, the radiologists received a re-randomized case collection for a second interpretation round according to the Fleischner and subsequently 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, and a definite or probable UIP pattern according to Fleischner. Inter-rater agreement between the three radiologists was assessed using Light-Kappa (κ), which is an extension of Cohen's Kappa for more than two raters. [7] Every radiological pattern was dichotomized (e.g. definite UIP YES/NO), and we considered a pattern to be present if 2-3 out of 3 radiologists scored the pattern with YES, otherwise the pattern was scored not to be present (NO). Test characteristics with 95% confidence intervals (95%CI) 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, NPV), and Youden's index which summarizes 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.

Results

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 were men (79%); 56 were ever-smokers (63%); the mean age was 68.4 years (standard deviation [SD] 10.4); mean body mass index 27.2 (SD 4.8) kg/m^2; mean forced vital capacity 68.4 (SD 19.0) %-predicted; mean diffusing capacity of the lung for carbon monoxide (DLCO) 50 (SD 16.7) %-predicted; with SLB available in 19 (33%) IPF and 21 (57%) patients with

other ILDs. During follow-up (mean 26.9 [interquartile range 17.2-50.2] months) 28 (54%) of all IPF and 2 (5.4%) of other ILD patients deceased.

The overall interrater agreement was good for ATS/ERS/JRS/ALAT2011 (κ 0.61), and moderate for ATS/ERS/JRS/ALAT2018 and Fleischner (κ 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 4/37 non-IPF cases as IPF, compared to 3/37 and 1/37 misclassifications with ATS/ERS/JRS/ALAT2018 and ATS/ERS/JRS/ALAT2011, respectively. Sensitivity was markedly higher for the Fleischner scenario with 7/52 missed IPF cases, compared to ATS/ERS/JRS/ALAT2018 and ATS/ERS/JRS/ALAT2011 with 20/52 and 23/52 missed IPF cases respectively. PPV and NPV was slightly higher for Fleischner than for ATS/ERS/JRS/ALAT2018, and of all scenarios, Fleischner had the strongest overall test performance (Table).

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). Fleischner scenario remained the strongest correlate of mortality risk including with adjustment for potential confounders (age, sex, ever smoker, and DLCO %-predicted [HR 3.76, 95%CI 1.22-11.6, model C-index 0.80]). Corresponding multivariate Cox regression models demonstrated the ATS2011 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).

Discussion

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 as well probable UIP pattern for IPF diagnosis in the right clinical context. In our cohort of well- characterized IPF patients compared to 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 und 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 definite HRCT UIP pattern. [8,9] In this cohort, including patients with a probable UIP pattern to the subgroup not requiring SLB in the diagnostic algorithm results in a significantly higher sensitivity with only marginal loss in specificity (incremental misclassification in 2.7%). Applying the Fleischner algorithm, we further found an increase in PPV and NPV compared to ATS/ERS/JRS/ALAT2018, with the limitation that these characteristics are not generalizable 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 inter-observer 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 non-invasive and if needed invasive biomarkers with inclusion of patient preference in the process.

References

1. Lynch DA, et al. Diagnostic criteria for idiopathic pulmonary fibrosis: a Fleischner Society White Paper. Lancet Respir Med. 2018 Feb;6(2):138-153.
2. Raghu G, et al; 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 Sep 1;198(5):e44-e68.
3. Richeldi L, Wilson KC, Raghu G. Diagnosing idiopathic pulmonary fibrosis in 2018: bridging recommendations made by experts serving different societies. Eur Respir J. 2018 Sep 6;52(3).
4. Wells AU. IPF diagnosis: flexibility is a virtue. Lancet Respir Med. 2018 Oct;6(10):735-737.
5. Raghu G, 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 Mar 15;183(6):788-824.
6. Guler S, Zumstein P, Berezowska S, Pöllinger A, Geiser T, Funke-Chambour M. Idiopathic pulmonary fibrosis in a Swiss interstitial lung disease reference centre. Swiss Med Wkly. 2018;148:w14577.
7. Light, R. J. (1971). Measures of response agreement for qualitative data: Some generalizations and alternatives. Psychological Bulletin, 76(5), 365-377.
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 May 29;370(22):2071-82.
9. Raghu G, Wells AU, Nicholson AG, Richeldi L, Flaherty KR, Le Maulf F, Stowasser S, Schlenker-Herceg R, Hansell DM. Effect of Nintedanib in Subgroups of Idiopathic Pulmonary Fibrosis by Diagnostic Criteria. Am J Respir Crit Care Med. 2017 Jan 1;195(1):78-85.
10. Brownell R, Moua T, Henry TS, Elicker BM, White D, Vittinghoff E, Jones KD, Urisman A, Aravena C, Johannson KA, Golden JA, King TE Jr, Wolters PJ, Collard HR, Ley B. The use of pretest probability increases the value of high-resolution CT in diagnosing usual interstitial pneumonia. Thorax. 2017 May;72(5):424-429.
11. Jo HE, Glaspole I, Goh N, Hopkins PMA, Moodley Y, Reynolds PN, Chapman S, Walters EH, Zappala C, Allan H, Macansh S, Grainge C, Keir GJ, Hayen A, Henderson D, Klebe S, Heinze SB, Miller A, Rouse HC, Duhig E, Cooper WA, Mahar AM, Ellis S, McCormack SR, Ng B, Godbolt DB, Corte TJ. Implications of the diagnostic criteria of idiopathic pulmonary fibrosis in clinical practice: Analysis from the Australian Idiopathic Pulmonary Fibrosis Registry. Respirology. 2019 Apr;24(4):361-368.
12. Walsh SL, Calandriello L, Sverzellati N, Wells AU, Hansell DM; UIP Observer Consort. Interobserver agreement for the ATS/ERS/JRS/LATS /JRS/ALAT criteria for a UIP pattern on CT. Thorax 2016 Jan;71(1):45-51. doi: 10.1136/thoraxjnl-2015-207252. Epub 2015 Nov 19.
13. Walsh SLF, Maher TM, Kolb M, Poletti V, Nusser R, Richeldi L, Vancheri C, Wilsher ML, Antoniou KM, Behr J, Bendstrup E, Brown K, Calandriello L, Corte TJ, Cottin V, Crestani B, Flaherty K, Glaspole I, Grutters J, Inoue Y, Kokosi M, Kondoh Y, Kouranos V, Kreuter M, Johannson K, Judge E, Ley B, Margaritopoulos G, Martinez FJ, Molina-Molina M, Morais A, Nunes H, Raghu G, Ryerson CJ, Selman M, Spagnolo P, Taniguchi H, Tomassetti S, Valeyre D, Wijsenbeek M, Wuyts W, Hansell D, Wells A; IPF Project Consortium. Diagnostic accuracy of a clinical diagnosis of idiopathic pulmonary fibrosis: an international case-cohort study. Eur Respir J. 2017 Aug 31;50(2).
14. Anthimopoulos M, Christodoulidis S, Ebner L, Geiser T, Christe A, Mougiakakou S.Semantic Segmentation of Pathological Lung Tissue with Dilated Fully Convolutional Networks. IEEE J Biomed Health Inform. 2019 Mar;23(2):714-722.

	Test characteristics*					Survival (unadjusted)		
Radiological scenario	**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**[†]	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**[†]	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[‡]	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

Table. Diagnostic and prognostic performance of radiological scenarios.

*Diagnostic test characteristics with IPF MDT diagnosis as the reference standard.
[†]Definite UIP scored according to ATS/ERS/JRS/ALAT2011 and *ATS/ERS/JRS/ALAT*2018 guidelines respectively
‡Definite and probable UIPs scored according to the recommendations by the Fleischner Society.

Abbreviations: CI, confidence interval; C-index, Harrell's concordance statistic; HR, hazard ratio; IPF, idiopathic pulmonary fibrosis; J-index, Youden's J statistic; NPV, negative predictive value; PPV, positive predictive value; UIP, usual interstitial pneumonia