Diagnostic accuracy of a clinical diagnosis of idiopathic pulmonary fibrosis: an international case–cohort study

Simon L.F. Walsh[1], Toby M. Maher[2], Martin Kolb[3], Venerino Poletti[4,5], Richard Nusser[6], Luca Richeldi[7], Carlo Vancheri[8], Margaret L. Wilsher[9], Katerina M. Antoniou[10], Jüergen Behr[11], Elisabeth Bendstrup[5], Kevin Brown[12], Lucio Calandriello[13], Tamera J. Corte[14], Vincent Cottin[15], Bruno Crestani[16], Kevin Flaherty[17], Ian Glaspole[18], Jan Grutters[19], Yoshikazu Inoue[20], Maria Kokosi[21], Yasuhiro Kondoh[22], Vasileios Kouranos[21], Michael Kreuter[23], Kerri Johannson[24], Eoin Judge[25], Brett Ley[26], George Margaritopoulos[21], Fernando J. Martinez[27], Maria Molina-Molina[28], António Morais[29], Hilario Nunes[30], Ganesh Raghu[31], Christopher J. Ryerson[32], Moises Selman[33], Paolo Spagnolo[34], Hiroyuki Taniguchi[22], Sara Tomassetti[4], Dominique Valeyre[30], Marlies Wijsenbeek[35], Wim Wuyts[36], David Hansell[37] and Athol Wells[38] on behalf of The IPF Project Consortium[39]

@ERSpublications
Academic status, access to MDT meetings and clinician experience predict accuracy of a clinical diagnosis of IPF http://ow.ly/k43W30cTMg1

Cite this article as: Walsh SLF, 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 [https://doi.org/10.1183/13993003.00936-2017].

ABSTRACT We conducted an international study of idiopathic pulmonary fibrosis (IPF) diagnosis among a large group of physicians and compared their diagnostic performance to a panel of IPF experts.

A total of 1141 respiratory physicians and 34 IPF experts participated. Participants evaluated 60 cases of interstitial lung disease (ILD) without interdisciplinary consultation. Diagnostic agreement was measured using the weighted kappa coefficient (κ_w). Prognostic discrimination between IPF and other ILDs was used to validate diagnostic accuracy for first-choice diagnoses of IPF and were compared using the C-index.

A total of 404 physicians completed the study. Agreement for IPF diagnosis was higher among expert physicians (κ_w=0.65, IQR 0.53–0.72, p<0.0001) than academic physicians (κ_w=0.56, IQR 0.45–0.65, p<0.0001) or physicians with access to multidisciplinary team (MDT) meetings (κ_w=0.54, IQR 0.45–0.64, p<0.0001). The prognostic accuracy of academic physicians with >20 years of experience (C-index=0.72, IQR 0.0–0.73, p=0.229) and non-university hospital physicians with more than 20 years of experience, attending weekly MDT meetings (C-index=0.72, IQR 0.70–0.72, p=0.052), did not differ significantly (p=0.229 and p=0.052 respectively) from the expert panel (C-index=0.74 IQR 0.72–0.75).

Experienced respiratory physicians at university-based institutions diagnose IPF with similar prognostic accuracy to IPF experts. Regular MDT meeting attendance improves the prognostic accuracy of experienced non-university practitioners to levels achieved by IPF experts.

This article has supplementary material available from erj.ersjournals.com

Received: May 07 2017 | Accepted after revision: May 30 2017

Conflict of interest: Disclosures can be found alongside this article at erj.ersjournals.com

Affiliations: [1]Dept of Radiology, King's College Hospital Foundation Trust, London, UK. [2]Dept of Respiratory Medicine, Interstitial Lung Disease Unit, Royal Brompton Hospital and National Heart and Lung Institute, Imperial College, London, UK. [3]Depts of Medicine and Pathology/Molecular Medicine, McMaster University, Firestone Institute for Respiratory Health, Hamilton, ON, Canada. [4]Department of Diseases of the Thorax, Ospedale GB Morgagni, Forlì, Italy. [5]Dept of Respiratory Diseases and Allergy, Aarhus University Hospital, Aarhus, Denmark. [6]Dept of Respiratory Medicine, Summit Hospital, Oakland, CA, USA. [7]Unità Operativa Complessa di Pneumologia, Università Cattolica del Sacro Cuore, Fondazione Policlinico A. Gemelli, Rome, Italy. [8]Dept of Clinical and Experimental Medicine, University of Catania, University - Hospital "Policlinico - Vitt. Emanuele", Catania, Italy. [9]Auckland District Health Board and the University of Auckland, Auckland, New Zealand. [10]Dept of Respiratory Medicine, Faculty of Medicine, University of Crete, Heraklion, Greece. [11]Dept of Medicine V, University of Munich and Asklepios Fachkliniken Gauting, Comprehensive Pneumology Center, member of the German Center for Lung Research, Munich, Germany. [12]National Jewish Health, Denver, CO, USA. [13]Institute of Radiology, Università Cattolica del Sacro Cuore, Fondazione Policlinico Universitario A. Gemelli, Rome, Italy. [14]Dept of Respiratory Medicine, Royal Prince Alfred Hospital, and University of Sydney, Sydney, Australia. [15]Service de Pneumologie, Hopital L. Pradel, Lyon, France. [16]APHP, Hopital Bichat, Service de Pneumologie A, Université Paris Diderot, Paris, France. [17]University of Michigan, Division of Pulmonary and Critical Care Medicine, Ann Arbor, MI, USA. [18]Alfred Health - Allergy, Immunology and Respiratory Medicine, Melbourne, Australia. [19]ILD Center of Excellence, St Antonius Hospital, Division Heart and Lungs, University Medical Centre Utrecht, Utrecht, The Netherlands. [20]Clinical Research Center, National Hospital Organization Kinki-Chuo Chest Medical Center, Osaka, Japan. [21]Dept of Respiratory Medicine, Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK. [22]Tosei General Hospital, Dept of Respiratory Medicine and Allergy, Seto, Japan. [23]Center for interstitial and rare lung diseases, Pneumology and respiratory critical care medicine, Thoraxklinik, University of Heidelberg, and Translational Lung Research Center Heidelberg, Member of the German Center for Lung Research (DZL), Heidelberg, Germany. [24]University Of Calgary, Calgary, AB, Canada. [25]Respiratory Medicine and National Lung Transplantation Unit, Mater Misericordiae University Hospital, Dublin , Ireland. [26]Medicine, University of California San Francisco, San Francisco, CA, USA. [27]Weill Cornell Medicine, New York, NY, USA. [28]University Hospital of Bellvitge-IDIBELL-CIBERES, Barcelona, Spain. [29]Centro Hospitalar São João – Pulmonology, Faculdade de Medicina do Porto, Alameda Professor Hernâni Monteiro, Oporto, Portugal. [30]Paris 13 University, Sorbonne Paris Cité, Service de Pneumologie, Hopital Avicenne, Bobigny, France. [31]University of Washington - Center for Interstitial Lung Disease, Seattle, WA, USA. [32]University of British Columbia – Medicine, Vancouver, BC, Canada. [33]Instituto Nacional de Enfermedades Respiratorias "Ismael Cosio Villegas", Mexico City, Mexico. [34]Section of Respiratory Diseases, Dept of Cardiac, Thoracic and Vascular Sciences, University of Padova, Padova, Italy. [35]Dept of Pulmonary Diseases, Erasmus MC, University Hospital Rotterdam, Rotterdam, The Netherlands. [36]Respiratory Medicine, University Hospitals Leuven, Leuven, Belgium. [37]Dept of Thoracic Imaging, Royal Brompton Hospital, London, UK. [38]Interstitial Lung Disease Unit, Royal Brompton Hospital, London, UK. [39]A full list of members of the IPF Project Consortium can be found in the acknowledgements section.

Correspondence: Simon L.F. Walsh, Kings College Hospital Foundation Trust, Denmark Hill, London SE5 9RS, UK. E-mail: slfwalsh@gmail.com

Introduction

Idiopathic pulmonary fibrosis (IPF) is characterised by progressive loss of lung function and a particularly poor prognosis [1]. Although it is often regarded as a rare disorder, in 2012 1% of all deaths in the UK occurred because of pulmonary fibrosis [2] and the incidence of IPF is expected to continue to rise [3–8]. Accurate IPF diagnosis has increased in importance with the advent of antifibrotic therapies and ongoing enrolment in IPF treatment trials [9, 10].

Although the ATS/ERS/JRS/ALAT guideline diagnostic recommendations emphasise the importance of a multidisciplinary approach when diagnosing IPF [11–13], less experienced non-academic clinicians outside regional centres may not have access to multidisciplinary team (MDT) meetings [14]. Therefore, the diagnostic accuracy of clinicians acting in isolation is of practical importance. In the absence of a reference standard, one approach to evaluating the diagnostic skills of clinicians is to examine separations in mortality between patients diagnosed with IPF and those diagnosed with other ILDs, a method used in a recent study of multidisciplinary diagnosis [15]. The most accurate discrimination between IPF and non-IPF diagnoses should, in principle, provide the greatest separation in outcomes.

The aim of this study was to evaluate and compare IPF diagnoses made by non-academic clinicians, university-affiliated clinicians and an international panel of IPF experts, using three surrogates of diagnostic accuracy: diagnostic confidence, diagnostic agreement and prognostic accuracy (which was examined in non-expert subgroups against years of experience and access to an MDT meeting).

Materials and methods

Case collection

The study protocol was approved by the NHS Health Research Authority, and for this retrospective examination of clinically indicated data, the need for patient consent was waived. We selected consecutive

patients presenting to the interstitial lung disease (ILD) unit of the Royal Brompton and Harefield NHS Foundation Trust (London, UK) between January 5, 2010, and October 25, 2010 (supplementary figure A1). This approach allowed an analysis of 5-year survival and also meant that patients included in the study were selected from a pre-antifibrotic therapy era. Therefore, outcome distinctions between patients with IPF and those without this disease were not confounded by antifibrotic therapy. Since referral rates of patients with suspected IPF to the host institution in 2010 differed (25% of all referrals) from 2015 (36% of all referrals), we enriched the cohort with consecutive patients referred to the host institution between January 5, 2010, and October 25, 2010, and who were diagnosed with IPF by the host institution, to match 2015 IPF referral rates. Exclusion criteria were as follows.

1) An established diagnosis of connective tissue disease prior to presentation to the host institution. In these patients, the diagnosis of connective tissue disease-related ILD is usually straightforward and might spuriously increase overall diagnostic agreement [15].

2) Non-availability of imaging or lung function tests at presentation.

3) Diffusing capacity of the lung for carbon monoxide (D_{LCO}) <30% predicted, excluded because: a) clinicians might assume that the presence of end-stage fibrosis indicates IPF thus impacting diagnostic agreement and accuracy for an IPF diagnosis; b) although patients with end-stage fibrotic lung disease may occasionally be referred to the host institution, this may not reflect referral patterns to less specialised centres; and c) treatment may be less effective in patients with end-stage fibrosis reducing the importance of diagnostic precision.

Participating physicians

Between January 1, 2015, and July 1, 2016, we performed an Internet search, country by country, for practising respiratory physicians. Physician experience, nationality, academic status (working at a university hospital or not a university hospital) or subspecialist interests within respiratory medicine did not influence inclusion eligibility. This search included the European Respiratory Society Diffuse Parenchymal Lung Disease Assembly and the American Thoracic Society Clinical Problems Assembly. During July 2016 an invitation to participate in the study was extended to all of the physicians identified. In addition to this group, an expert panel was created, comprising respiratory physicians with specialist expertise in the diagnosis and management of ILD working in specialist ILD centres and with a track record of publications in this field. For the purposes of this study, physicians working at university-affiliated institutions are referred to as university physicians or academic physicians, and physicians not working at university-affiliated institutions are referred to as non-university physicians or non-academic physicians.

Scoring protocol

Evaluation of cases took place between July 1, 2016, and January 1, 2017, on a custom built web-based application. First, physicians were required to answer a preliminary survey regarding their usual clinical practice (supplementary table A1). Then for each case they were presented with the patient's history, findings on physical examination and standardised baseline clinical information, extracted from the patient electronic records (supplementary table A2). Physicians were provided the presentation high-resolution computed tomography scan (HRCT). The original HRCT report was not provided. We did not inform physicians if the host institution had performed surgical lung biopsy. Since biopsy decisions depend on a physician's individual clinical judgement, there would be no way of knowing which patients would eventually have undergone a lung biopsy. Also, if we had provided biopsy information, the clinical skill of the physician would be influenced by the expertise of the host institution.

The scoring protocol has been described previously [15]. For each case, physicians were required to select up to five differential diagnoses and provide a diagnostic likelihood (censored at 5% and summing to 100% in each case) from a drop-down menu of diffuse lung diseases (supplementary table A3) based upon their diagnostic confidence. The drop-down menu included a category labelled "other", to be selected when the desired diagnosis was not listed. In this situation, physicians were required to provide their diagnosis in a free-text box. The only stipulation to scoring the cases was that each case was evaluated in isolation without interspecialty consultation.

Statistical analysis

Statistical analyses were performed using STATA (version 14, StataCorp, College Station, TX, USA). Data are given as means with standard deviations (SD), medians with interquartile range (IQR) or as the number of patients and percentage where appropriate. Group comparisons were made using the *t*-test, Wilcoxon rank sum, chi-squared statistics and Fisher's exact test where appropriate.

Cohen's kappa coefficient (κ) was used to evaluate interobserver agreement for diagnosis and the weighted kappa coefficient (κ_w) was used to evaluate interobserver agreement for an estimation of the probability of each diagnosis. To do this, the percentage diagnostic likelihood given for each diagnosis was converted to a 5-point scale (0–4), representing clinically useful probabilities: 0=condition not included in the differential diagnosis, 1=low probability (5–25%), 2=intermediate probability (30–65%), 3=high probability (70–95%) and 4=pathognomonic (100%). This approach has been used in previous investigations of interobserver agreement for the diagnosis of diffuse lung diseases (supplementary material) [15–17]. Weighted kappa values were calculated between paired observers and expressed as median values with interquartile ranges for all unique combinations of pairs. Weighting the kappa coefficient allowed the degree of disagreement to be quantified by assigning greater emphasis to large differences between scores. Additionally, for each patient the first-choice diagnosis was considered high confidence if the diagnostic likelihood assigned was ⩾70%. This distinction is based on the diagnostic likelihood categories used to assess the clinical probability of pulmonary embolism in the PIOPED study [18] and has been used in another study of diagnostic agreement [15].

We used outcome distinctions between IPF and other diffuse lung diseases to validate diagnostic accuracy for IPF by converting each physician's first-choice diagnosis into a binary IPF diagnosis category (IPF or not IPF). Cox proportional hazards modelling was then used to determine a hazard ratio (HR) for each physician, adjusted for disease severity by including per cent predicted *D*LCO in the regression model. Time to death was the outcome for survival analyses and the survival period for each patient was calculated from the date of referral to the host institution to January 1, 2015. We tested the assumptions of proportional hazards by visual inspection of the log–log plot of survival, comparison of the Kaplan–Meier observed survival curves with the Cox predicted curves for the same variable and graphical and formal analysis of Schoenfeld residuals (analysis not shown). Results are reported as HR, 95% CI and p values. The prognostic accuracy of individual physician diagnoses was quantified using Harrell's C-index, which when used in this context, is a measure of prognostic discrimination (supplementary material) [19].

Multivariate linear regression models were used to identify independent predictors of prognostic accuracy within physician subgroups (expert, university, non-university and those with and without access to MDT meetings) taking the C-index as the dependent variable and using a backward elimination procedure, retaining variables with p<0·05. The assumptions of linear regression were tested and confirmed by inspection of residual-*versus*-predictor plots and heteroskedasticity was tested for graphically (by inspection of residuals plotted against fitted values) and non-graphically (using the Cook–Weisberg test for heteroskedasticity). The diagnostic performance of various subgroups of physicians based on these predictors was then compared to the expert panel group

Results

Patient population and participating physicians

The total cohort of cases was made up of 60 patients, including 22 (36.7%) with an MDT meeting diagnosis of IPF. Five patients required surgical lung biopsy. Three of these were diagnosed as IPF, one as pulmonary alveolar proteinosis and one as obliterative bronchiolitis. Vital status was known for all patients at the end of the study period. There were 26 out of 60 (43.4%) deaths at the end of the study period. Mean follow-up periods for IPF and non-IPF cases were 1246.0 days and 1646.0 days respectively. For more details of patient exclusions, diagnoses and mortality, see the supplementary material and supplementary table A4.

A total of 1141 respiratory physicians from 102 countries were invited to participate in the study. Between July 7, 2016, and January 1, 2017, 750 physicians representing 76 countries enrolled and completed the preliminary survey. Of these, 404 physicians, representing 57 countries, which included a panel of 34 invited experts, completed the evaluation of all 60 cases. Physicians who completed the study were more likely to be fellowship trained, work at university hospitals, have access to MDT meetings and diagnose more cases of IPF per month (tables 1 and 2). A summary of physician demographics based on country is shown in supplementary table A5.

Frequency of IPF diagnosis and diagnostic confidence

A total of 24 240 case evaluations were performed (404 physicians×60 cases). IPF made up 6308 (26.0%) of all first-choice diagnoses. Of the IPF diagnoses, 72.3% were made with high confidence (diagnostic likelihood ⩾70%). Expert panel members and academic physicians made high confidence diagnoses of IPF more frequently than non-academic physicians (p=0.002 and p=0.001, table 3) and more frequently diagnosed IPF overall (p=0.005 and p=0.008, table 3). Attendance at MDT meetings was not associated with a higher frequency of IPF diagnoses or a higher frequency of highly confident IPF diagnoses (p=0.718, p=0.925, table 3).

Diagnostic agreement

Overall interobserver agreement for the first-choice diagnosis of IPF was moderate for the entire cohort of physicians (n=404, κ=0.42). Unweighted Kappa values for interobserver agreement for a diagnosis of IPF for various physician subgroups are shown in table 4. The greatest diagnostic agreement for the first-choice diagnosis of IPF was between the expert panel members (n=34, κ=0.53). Physicians with no access to MDT meetings had the lowest level of diagnostic agreement for the first-choice diagnosis of IPF (n=76, κ=0.35) (table 4). Agreement on the likelihood of an IPF diagnosis (ranging from <5% to >95%) was highest among expert physicians, academic physicians and physicians with access to MDT meetings (table 5). Interobserver agreement for the likelihood of an IPF diagnosis between physicians based on country is shown in supplementary table A6.

Prognostic accuracy of an IPF diagnosis

Diagnoses of IPF were prognostically significant for 318 of 404 respiratory physicians (68.6%, median HR 2.81, IQR 2.21–3.61; median C-index=0.72, IQR 0.70–0.74). The range of C-indices across the entire cohort of physicians was 0.69–0.81. Hazards ratios, p values and C-indices for all participating physicians based on country are shown in supplementary table A7. Expert physicians, compared to other physicians, were more likely to make prognostically significant IPF diagnoses (29/34, 85.2%, *versus* 246/370, 66.4%; p=0.02) and with greater prognostic discrimination (as judged by C-indices), p=0.0002 (supplementary table A8). Academic physicians demonstrated greater prognostic discrimination for a diagnosis of IPF than non-university based hospital physicians, p=0.0006 (supplementary table A9). Physicians who attend MDT meetings demonstrated greater prognostic discrimination for a diagnosis of IPF than physicians not attending MDT meetings, p=0.004, supplementary table A10).

Multivariate linear regression analysis was performed taking the C-index as the dependent variable and 1) academic status, 2) years of experience (stratified by thresholds ranging from 5–35 years in 5-year

TABLE 1 Responses to the preliminary survey by 404 physicians who completed the study and the 346 physicians who did not complete the study

Question	Completed (n=404)	Did not complete (n=346)	p-value
Experience years	15.8	15.7	0.565
ILD Fellowship training			
Yes	359 (88.9%)	283 (70.0%)	0.006
In-training	17 (4.2%)	20 (5.8%)	0.322
No	28 (6.9%)	43 (12.4%)	
Hospital setting			
University	288 (71.3%)	207 (59.8%)	0.001
Not university	116 (28.7%)	139 (40.2%)	
MDT meeting			
MDT meeting access	328 (81.2%)	247 (61.8%)	0.002
No MDT meeting access	76 (18.8%)	99 (28.6%)	
Number of cases of IPF diagnosed/month			
None, we refer all cases of suspected IPF to an academic centre	20 (5.0%)	38 (11.0%)	0.002
1–10	337 (83.4%)	290 (83.8%)	0.883
11–20	37 (9.2%)	12 (3.5%)	0.002
20+	9 (2.2%)	5 (4.1%)	0.430
Access to specialist radiology expertise			
None	22 (5.4%)	26 (7.5%)	0.191
Not directly but in my network	60 (14.9%)	58 (16.8%)	0.474
Yes	322 (79.7%)	262 (75.7%)	0.248
Access to specialist pathology expertise			
None	34 (8.4%)	35 (10.1%)	0.006
Not directly but in my network	85 (21.0%)	100 (28.9%)	0.013
Yes	285 (70.5%)	211 (61.0%)	0.422
Availability of cryobiopsy			
Yes	65 (16.1%)	44 (12.7%)	0.191
No	339 (83.9%)	302 (87.3%)	

ILD: interstitial lung disease; MDT: multidisciplinary team; IPF: idiopathic pulmonary fibrosis.

TABLE 2 Responses to the preliminary survey by 404 physicians grouped according to institution type (university hospital or not university hospital)

Grouping	University hospital (n=288)	Not university hospital (n=116)	p-value
Experience years	14.9	17.8	0.009
Fellowship trained	251	108	0.085
MDT meeting practices			
No MDT meeting	41	35	0.001
Daily MDT meeting	4	0	0.202
Weekly MDT meeting	118	25	0.001
Fortnightly MDT meeting	41	8	0.021
Monthly MDT meeting	66	34	0.178
Less than 1/month MDT meeting	18	14	0.05
Number of IPF cases diagnosed/month			
Refer all cases of suspected IPF	14	6	0.896
1–10 cases	234	103	0.065
11–20 cases	32	5	0.032
20+ cases	7	2	0.663
Access to radiology expertise			
Direct access	242	80	0.001
Access through network	35	25	0.016
No access	11	11	0.023
Access to pathology expertise			
Direct access	219	66	0.001
Access through network	49	36	0.002
No access	20	14	0.093
Cryobiopsy part of usual practice	54	11	0.022

MDT: multidisciplinary team; IPF: idiopathic pulmonary fibrosis.

increments), 3) MDT meeting attendance and 4) the number of IPF cases diagnosed per month as the independent variables. Academic status, >20 years of experience and attendance at MDT meetings independently predicted the prognostic accuracy of IPF diagnosis (supplementary table A11). Subsequent

TABLE 3 Median number of idiopathic pulmonary fibrosis (IPF) diagnoses made and median number of high confidence IPF diagnoses made for individual physicians by subgroup

	Expert panel physicians (n=34)	Others (n=370)	p-value
Median number of IPF diagnoses	20 (IQR 14–23)	15 (IQR 11–19)	0.005
Median number of high confidence IPF diagnoses	17 (IQR 8–21)	11 (IQR 7–14)	0.002
	University hospital physicians (n=288)	**Not university hospital physicians (n=116)**	**p-value**
Median number of IPF diagnoses	16 (IQR 12–20)	13 (IQR 10–19)	0.008
Median number of high confidence IPF diagnoses	11 (IQR 8–16)	9 (IQR 6–12)	0.001
	MDT meeting attendance (n=328)	**No MDT meeting attendance (n=76)**	**p-value**
Median number of IPF diagnoses	15 (IQR 11–20)	15 (IQR 10–20)	0.925
Median number of high confidence IPF diagnoses	11 (IQR 7–15)	11 (IQR 6.5–15)	0.718

All values are out of 60 cases. High confidence diagnoses are defined as those cases assigned a diagnosis of IPF with a diagnostic likelihood of ⩾70%. MDT: multidisciplinary team; IQR: interquartile range.

analyses of particular interest are summarised in table 6. Specifically, 1) university hospital physicians with >20 years of experience achieved equivalent prognostic discrimination to the expert panel for a diagnosis of IPF (or not IPF group), regardless of attendance at weekly MDT meetings (table 6, supplementary figure A2); 2) non-university hospital physician prognostic discrimination did not reach that of the expert panel, regardless of availability of MDT meetings or the threshold of 20 years of experience (table 6, supplementary figure A2); 3) however, non-university hospital physicians with >20 years of experience, attending *weekly* MDT meetings, demonstrated near expert level prognostic accuracy (C-index 0.72, IQR 0.70–0.72; p=0.052).

Discussion

Our results show that academic status, attendance at MDT meetings and experience level of physicians are independently associated with greater prognostic discrimination between diagnoses of IPF and other ILDs. In particular, using mortality to validate accuracy of IPF diagnosis, we have shown that accuracy of IPF diagnosis made by university hospital-based practitioners with greater than 20 years of experience is equivalent to that of international IPF experts.

A recent study reported near parity in diagnostic agreement and accuracy for IPF between expert physicians and their respective MDT meetings [15]. The purpose of this study was to investigate whether these findings could also be applied to physicians of varying levels of experience when acting in isolation without the benefit of MDT meeting evaluation. A central feature of our study was that we validated IPF diagnosis against mortality, an approach used in a previous study of diagnostic agreement and accuracy in IPF [15]. In diffuse lung disease, multidisciplinary discussion is the recommended approach to diagnosis, which involves integrating all available clinical, radiologic and if available, pathologic data. For this reason, there is no reference standard against which the veracity of MDT diagnosis can be tested. However, as a poor outcome is a cardinal feature of IPF, accurate diagnosis should, in principle, provide the greatest prognostic discrimination between IPF and other ILDs.

The use of the C-index to examine prognostic accuracy between physician subgroups warrants further discussion. The range of achievable C indices for the 404 clinicians was narrow (from 0.69 to 0.81, therefore representing a 13-point scale) and is likely to reflect the fact that some non-IPF patients do badly and so misclassification of a non-IPF patient as an IPF patient might not significantly impact the C-index. Given this outcome overlap between the two disease groups, apparently small cohort shifts are likely to be more meaningful than they appear. Furthermore, as an example, the difference between a C-index of 0.70 and 0.72 is 15.4%, representing an upward shift in prognostic accuracy of 15.4% when comparing a population of experienced academic clinicians with MDT access to a population of less experienced (<20 years) non-academic clinicians without MDT access. This is in fact a large difference when, for example, it is compared with cohort shifts in serial forced vital capacity (FVC) in IPF antifibrotic trials, put up against the range of baseline FVC values.

Although several studies have reported that MDT diagnosis is associated with higher levels of diagnostic confidence and superior interobserver agreement when compared with the individual components of the MDT in isolation [15, 20, 21], the effect that MDT meetings have on individuals has not been examined. One of the assumed benefits of a multidisciplinary approach to IPF diagnosis is that those participating have their diagnostic thinking subjected to peer scrutiny. The regular interspecialty discussion that MDT meetings promote is likely to broaden a physician's experience and establish an ethos of debate and critical evaluation. Conceivably, physicians who are accustomed to this process gain skill in related disciplines such as HRCT interpretation, which they can use outside the multidisciplinary setting. For some

TABLE 4 Unweighted kappa values (κ) for interobserver agreement for a diagnosis of idiopathic pulmonary fibrosis for various physician subgroups

Group	Interobserver agreement (κ)
Physicians, expert panel (n=34)	0.53
Physicians, non expert panel (n=370)	0.41
University physicians (n=288)	0.43
Non-university physicians (n=116)	0.38
Physicians with MDT meeting access (n=328)	0.44
Physicians without MDT meeting access (n=76)	0.35

MDT: multidisciplinary team.

TABLE 5 Comparisons of weighted kappa values (κ_w) for interobserver agreement on the diagnostic likelihood of a diagnosis of idiopathic pulmonary fibrosis between various subgroups

Group comparisons	Interobserver agreement (κ_w)	p
Physicians, expert panel (n=34)	0.65 (IQR 0.53–0.72)	<0.001
Remaining physician group (n=370)	0.53 (IQR 0.41–0.63)	
University hospital physicians (n=288)	0.56 (IQR 0.45–0.65)	<0.001
Non-university hospital physicians (n=116)	0.49 (IQR 0.38–0.59)	
MDT meeting available (n=328)	0.56 (IQR 0.45–0.65)	<0.001
No MDT meeting available (n=76)	0.46 (IQR 0.33–0.58)	

MDT: multidisciplinary team.

physicians, increasing patient numbers and possibly referrals from other centres will mean that full MDT meeting characterisation is possible only for selected cases. Therefore, just as in this study, it is likely that a substantial number of IPF patients will receive a diagnosis made by their respiratory physician acting in isolation. In a recent national survey conducted in France, IPF diagnosis resulted from multidisciplinary discussion in only 50% of cases [22]. It is noteworthy that in the current study, 43% of completing physicians stated that in most cases of suspected IPF, they made the diagnosis by themselves with the aid of diagnostic guidelines. We demonstrate that weekly MDT meeting attendance among experienced non-university hospital physicians increased prognostic accuracy of IPF diagnosis to that achieved by IPF experts.

Our findings may have implications for future multidisciplinary practice. Based on several studies of diagnostic agreement and accuracy over the past decade, MDT evaluation of IPF has become enshrined in the literature as the optimum approach to diagnostic synthesis [1, 12, 13, 15, 20, 23, 24]. A difficulty implementing this recommendation is that local access to multidisciplinary expertise may be limited. One possible solution to this problem is to network with academic centres using different forms of telemedicine. Since the web-based evaluation of patients in this study to some extent replicates telemedicine methodologies, our findings provide support for telemedicine as an acceptable form of multidisciplinary practice [24, 25]. Such collaboration could also include guidance on setting up local community hospital MDT meetings or having community physicians attend MDT meetings at local university hospitals.

Our study has some unavoidable limitations, common to previous studies of multidisciplinary practice [15–17]. First, unlike real-world clinical practice, it was impractical for physicians to engage in face-to-face consultation with patients, meaning that doctors did not have the chance to take a clinical history or examine the patients themselves. In complex disease, direct contact with the patient may influence a clinician's impression in a manner that is not easy to quantify objectively. However, direct patient contact in a study of this size would have been impracticable. Our methodology of web-based case reviews is instead similar to that of previously published studies of diagnostic agreement and accuracy between MDT meetings [15, 20, 26]. Second, physicians who completed the study were more likely to be fellowship

TABLE 6 Prognostic accuracy expressed as the C-index for diagnosis of idiopathic pulmonary fibrosis (IPF) or not IPF given by various physician subgroups.

University hospital physicians			Non-university hospital physicians		
Group	**C-index**	**p-value**	**Group**	**C-index**	**p-value**
>20 years of experience, no MDT meeting (n=11)	0.72 (0.70–0.73)	0.229	>20 years of experience, no MDT meeting (n=18)	0.70 (0.70–0.73)	0.008
>20 years of experience, MDT meeting (n=51)	0.72 (0.71–0.75)	0.116	>20 years of experience, MDT meeting (n=24)	0.71 (0.70–0.73)	0.019
<20 years of experience, no MDT meeting (n=30)	0.71 (0.70–0.72)	<0.001	<20 years of experience, no MDT meeting (n=17)	0.70 (0.70–0.71)	<0.001
<20 years of experience, MDT meeting (n=167)	0.72 (0.70–0.74)	0.001	<20 years of experience, MDT meeting (n=52)	0.71 (0.69–0.72)	<0.001

p-values are based upon a group comparison with the expert panel (n=34, C-index=0.74 (0.72–0.75)). MDT=multidisciplinary team.

trained, work at university institutions, attend MDT meetings and diagnose more cases of IPF per month. Nevertheless, sufficient numbers of physicians working in non-university institutions and without access to MDT meetings took part in our study, allowing us to perform statistically meaningful analyses in these subgroups. Third, to our knowledge no guideline recommendation indicates what precisely constitutes a valid MDT meeting. Although we asked physicians if they participated in formal MDT meetings, we did not attempt to quantify informal interspecialty consultation, which might also be considered by some to be a form of multidisciplinary practice [25]. An investigation to identify the optimum MDT meeting format could be the focus of future investigations. Fourth, although poor outcome separates IPF from other ILDs, the natural history of IPF is heterogeneous and therefore there is likely to be some overlap between the two disease groupings (IPF *versus* not IPF). Finally, although our selection of cases from 2010 meant that mortality differences between patients with IPF and other ILDs were not confounded by treatment with antifibrotic therapy, we did not evaluate the potential confounding influences of immunosuppressive therapy (which may be harmful in IPF patients) on mortality.

In conclusion, our study indicates that diagnostic agreement for IPF is acceptable among a large group of respiratory physicians of varying degrees of experience and drawn from a wide range of geographic locations. However, experienced respiratory doctors who work at university-based institutions show greater agreement on a diagnosis of IPF and make greater prognostic distinctions between IPF and other diffuse lung diseases than those at non-university institutions. Importantly, the diagnostic performance of experienced non-university practitioners improves with regular MDT meetings. These results may be a stimulus for greater interaction between university and community hospitals as well as the development of local MDT meetings for the specific purpose of assessing patients with suspected IPF.

Acknowledgements

Members of the IPF Project Consortium are as follows: Ping Shi Zhu, Yiming Yuan, Cesar Yoshito Fukuda, Yuki Yoshimatsu, Antoni Xaubet, Ai Ming Wong, Peter White, Gloria Westney, Alex West, Thomas Wessendorf, Yuko Waseda, Chingyu Wang, Jose Maria Vienna, Jelica Videnovic Ivanov, Vanesa Vicens Zygmunt, Maria Del Carmen Venero Caceres, Gerson Velásquez Pinto, Elizabeth Veitch, Martina Vasakova, Francesco Varone, Brenda Elena Varela, Peter Van Hal, Marjo Van De Ven, Ivo Van Der Lee, Leon Van Den Toorn, Amaia Urrutia Gajate, Ján Urban, Luis Gonzalo Ugarte Fornell, Argyrios Tzouvelekis, Kevin Twohig, Alice Turner, Sergio Trujillo, Antonio Triani, Daniel Traila, Veronica Torres, Hiromi Tomioka, Keisuke Tomii, Rade Tomic, Claudia Toma, Fatma Tokgoz Akyil, Kazunori Tobino, Ricardo Tobar, Anupama Tiwari, Regina Tibana, Xinlun Tian, Muhunthan Thillai, Weiping Tham, Felicia Teo, Jasna Tekavec Trkanjec, Paulo Teixeira, David Tarpey, Leonidas Tapias, Kiminobu Tanizawa, Yoshinori Tanino, Toshinori Takada, Gabriela Tabaj, Erzsebet Szolnoki, Rajesh Swarnakar, Irina Strambu, Martina Sterclova, Katherine Spinks, Chun Ian Soo, Amir Soltani, Sonam Solanki, Eman Sobh, Maria Raquel Soares, Jan Smith, Brian Smith, Philip Slocum, Hans Slabbynck, Ilya Sivokozov, Adrian Shifren, Shih Min Shen, Charles Sharp, Aruna Shanmuganathan, Alfredo Sebastiani, Simone Scarlata, Recep Savas, Shinichi Sasaki, Joanna Santeliz, Alfredo N.C. Santana, Rafaela Sanchez, Mauricio Salinas, Shigeki Saito, Frank Ryan, Juan Antonio Royo Prats, Elisabetta Rosi, Haala Rokadia, Alejandro Robles Perez, Pilar Rivera Ortega, Maite Rio Ramirez, Stefano Righetti, Cristina Reichner, Claudia Ravaglia, Pailin Ratanawatkul, Vijaya Ramalingam, Arvind Rajasekaran, Elzbieta Radzikowska, Seung Won Ra, Silvia Quadrelli, Juan Precerutti, Jyotika Prasad, Daniel Popa, Silvia Pizzolato, Wojciech Piotrowski, Andres Pineiro, Davide Piloni, Tatjana Peroš Golubičić, Rafael Perez, Carlos Pereira, Brett Pereira, Michael Perch, Nina Patel, Dharmesh Patel, Ilias Papanikolaou, Despina Papakosta, Efstratios Panselinas, Yong Kek Pang, Priyada Pandya, Eva Padrão, Ozlem Ozdemir Kumbasar, Maria J. Overbeek, Armine Otto Minasian, Deirdre O'Riordan, Josuel Ora, Justin Oldham, Oguzhan Okutan, Shinichiro Ohshimo, I.Kivilcim Oguzulgen, Takashi Ogura, Timothy O'Donnell, Conor O'Dochartaigh, Sarah O'Beirne, Mario Nunes, Lubov Novikova, Luca Novelli, Imre Noth, Nilson Nogueira Mendes Neto, Mitra Niroumand, Asunción Nieto, Alexandra Neves, Anoop Nambiar, Sudhir Nair, Rufai Nadama, Eoin Murtagh, Marco Mura, Joachim Müller Quernheim, Ansuman Mukhopadhyay, Suranjan Mukherjee, Julie Morisset, Onofre Moran, Joshua Mooney, Janne Møller, Nesrin Mogulkoc, Atsushi Miyamoto, Branislava Milenkovic, Stephen Mette, Mayra Mejia, Federico Mei, Mariano Mazzei, Toshiaki Matsuda, Carol Mason, Manuela Martínez Francés, George Margaritopoulos, Silvina Mannarino, Eliane Mancuzo, Foteini Malli, Puneet Malhotra, Martin Maillo, Jose Maia, Masoud Mahdavian, Flemming Madsen, Tracy Luckhardt, William Lucht, Su Ying Low, Concepcion Patricia Lopez Miguel, Randolph Lipchik, Steven Levy, Kovi Levin, Kai Ling Lee, David Lederer, Matthew Lammi, Hoi Yee Kwan, Sandeep Kukreja, Anuk Kruavit, Vasilis Kouranos, Mateusz Kotecki, Likurgos Kolilekas, Maria Kokosi, Heiko Knoop, Esen Kiyan, Tomoo Kishaba, Melissa King Biggs, Yet H Khor, Ayaz Khan, Nasreen Khalil, Ravindra Kedia, Naomi Kebba, Leticia Kawano Dourado, Kent Kapitan, Chung Dann Kan, Ali Fuat Kalyoncu, Meena Kalluri, Yasemin Kabasakal, Soma Jyothula, Maria Antonia Juretschke, Eoin Judge, Dragana Jovanovic, René Jonkers, Helen Jo, Shinyu Izumi, Haruyuki Ishii, Satoshi Ikeda, Aishah Ibrahim, Abdalla Ibrahim, Arfian Ibrahim, Charlotte Hyldgaard, Gary Hunninghake, Tristan Huie, Alison Hufton, Xiaowen Hu, Wan Chin Hsieh, Rodrigo Hoyos, Rachel Hoyles, Omar Holguin Rodriguez, Marie Pier Hogan, Ulla Hodgson, Helen Hilkin Sogoloff, Edwin Herrera, Brandon Michael Henry, Merel Hellemons, Ana Hecimovic, Ryuji Hayashi, Simon Hart, Sergio Harari, Sarah Haney, Nathan Hambly, Rismon Hakkim, Mónica Gutiérrez, Ria Gripaldo, Alejandro Gomez, Nicole Goh, Raul Godoy, Christopher Gilbert, Ioannis Giannarakis, Stefano Gasparini, Puneet Garcha, Sofia Furtado, Alessandro Fois, Patrick Flood Page, Sophie Fletcher, Elie Fiss, Juan Figueroa Casas, Marcelo Figueroa Casas, Christine A. Fiddler, Giovanni Ferrara, Marcelo Fernandez Casares, Charles Felton, Paola Faverio, Alexandre Todorovic Fabro, Andrea Estrada, Peter Errhalt, Noriyuki Enomoto, Juan Ignacio Enghelmayer, Karim El Kersh, Glenn Eiger, Anna Dubaniewicz, Fotios Drakopanagiotakis, Supparerk Disayabutr, Antoon Dijkstra, Juan Carlos Diaz Patiño, Jesus Javier Diaz Castañon, Sahajal Dhooria, Devesh J Dhasmana, Marino De Rosa , Stefano De Luca, Alain Delobbe, Diemen Delgado, Cesar Delgado, Isabel De La Fuente, Martijn De Kruif, Martin De Gier, Joao De Andrade, Jesper Rømhild Davidsen, Baseem Daoud, Klaus Dalhoff, Juan Vicente Cotera Solano, Andre Nathan Costa, Soledad Coronel, Marco

Confalonieri, Lennart Conemans, Alejandro Comellas, Sara Colella, Susana Clemente, John Clark, Marcello Ciuffreda, Chi Li Chung, Sy Giin Chong, Daniela Chirita, Po Lin Chen, Nazia Chaudhuri, Daniel Chambers, George Chalmers, Dennis Chairman, Gin Tsen Chai, Ronald Chacon Chaves, Veli Çetinsu, Michele Ceruti, Cecilio Omar Ceballos Zuñiga, Diego Castillo, Roberto G Carbone, Antonella Caminati, Francisco Javier Callejas González, Marcus Butler, Cesar Bustos, Marcin Bukowczan, Ivette Buendia, Giuseppe Brunetti, Angel Brown, Ben Brockway, Paul Bresser, José Breseghello, Demosthenes Bouros, Juan Antonio Botero Zaccour, Gisella Borzone, Raphael Borie, Hans Christian Blum, Jacky Blank, Abhishek Biswas, David Bennett, Mark Benjamin, Ionela Nicoleta Belaconi, Paul Beirne, Lutz Beckert, Shalini Bastiampillai, Rebecca Bascom, Brian Bartholmai, Manuel Barros, Andrea Yu Lin Ban, Elisabetta Balestro, Bruno Baldi, José Baddini Martinez, Archana Baburao, Suresh Babu, Alexander Averyanov, Sergey Avdeev, Rodrigo Athanazio, Ersan Atahan, Bassey Asuquo, Deborah Assayag, Julio Antuni, Sofia Antillon, Kenneth C Anderson, Adam Anderson, Fida Alwani, Goksel Altinisik, Neamat Alsouofi, J. Shirine Allam, Hamdan Al Jahdali, Abdulla Al Farttoosi, Tiago Alfaro, Nasser Al Busaidi, Ali Alavi Foumani, Maria Gloria Agreda Vedia, Amit Agarwal, Faraz Afridi, Olufunke Olayinka Adeyeye, Ayodeji Adegunsoye, Huzaifa Adamali, Atefeh Abedini.

References

1 Raghu G, Collard HR, Egan JJ, *et al.* 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.
2 British Lung Foundation. Idiopathic pulmonary fibrosis statistics. 2016.
3 Gribbin J, Hubbard RB, Le Jeune I, *et al.* Incidence and mortality of idiopathic pulmonary fibrosis and sarcoidosis in the UK. *Thorax* 2006; 61: 980–985.
4 Hutchinson J, Fogarty A, Hubbard R, *et al.* Global incidence and mortality of idiopathic pulmonary fibrosis: a systematic review. *Eur Respir J* 2015; 46: 795–806.
5 Hutchinson JP, McKeever TM, Fogarty AW, *et al.* Increasing global mortality from idiopathic pulmonary fibrosis in the twenty-first century. *Ann Am Thorac Soc* 2014; 11: 1176–1185.
6 Navaratnam V, Fleming KM, West J, *et al.* The rising incidence of idiopathic pulmonary fibrosis in the U.K. *Thorax* 2011; 66: 462–467.
7 Navaratnam V, Fogarty AW, Glendening R, *et al.* The increasing secondary care burden of idiopathic pulmonary fibrosis: hospital admission trends in England from 1998 to 2010. *Chest* 2013; 143: 1078–1084.
8 Olson AL, Swigris JJ. Idiopathic pulmonary fibrosis: diagnosis and epidemiology. *Clin Chest Med* 2012; 33: 41–50.
9 Richeldi L, du Bois RM, Raghu G, *et al.* Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2071–2082.
10 King TE Jr, Bradford WZ, Castro-Bernardini S, *et al.* A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014; 370: 2083–2092.
11 American Thoracic Society/European Respiratory Society International Multidisciplinary Consensus Classification of the Idiopathic Interstitial Pneumonias. *Am J Respir Crit Care Med* 2002; 165: 277–304.
12 Travis WD, Costabel U, Hansell DM, *et al.* An official American Thoracic Society/European Respiratory Society statement: Update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med* 2013; 188: 733–748.
13 Raghu G, Rochwerg B, Zhang Y, *et al.* An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *Am J Respir Crit Care Med* 2015; 192: e3–19.
14 Cottin V, Cadranel J, Crestani B, *et al.* Management of idiopathic pulmonary fibrosis in France: a survey of 1244 pulmonologists. *Respir Med* 2014; 108: 195–202.
15 Walsh SL, Wells AU, Desai SR, *et al.* Multicentre evaluation of multidisciplinary team meeting agreement on diagnosis in diffuse parenchymal lung disease: a case-cohort study. *Lancet Respir Med* 2016; 4: 557–565.
16 Aziz ZA, Wells AU, Hansell DM, *et al.* HRCT diagnosis of diffuse parenchymal lung disease: inter-observer variation. *Thorax* 2004; 59: 506–511.
17 Walsh SL, Calandriello L, Sverzellati N, *et al.* Interobserver agreement for the ATS/ERS/JRS/ALAT criteria for a UIP pattern on CT. *Thorax* 2016; 71: 45–51.
18 Pioped Investigators. Value of the ventilation/perfusion scan in acute pulmonary embolism. Results of the prospective investigation of pulmonary embolism diagnosis (PIOPED). *JAMA* 1990; 263: 2753–2759.
19 Harrell FE Jr, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med* 1996; 15: 361–387.
20 Flaherty KR, King TE Jr, Raghu G, *et al.* Idiopathic interstitial pneumonia: what is the effect of a multidisciplinary approach to diagnosis? *Am J Respir Crit Care Med* 2004; 170: 904–910.
21 Thomeer M, Demedts M, Behr J, *et al.* Multidisciplinary interobserver agreement in the diagnosis of idiopathic pulmonary fibrosis. *Eur Respir J* 2008; 31: 585–591.
22 Cottin V, Bergot E, Bourdin A, *et al.* Adherence to guidelines in idiopathic pulmonary fibrosis: a follow-up national survey. *ERJ Open Res* 2015; 1: 00032-2015.
23 Flaherty KR, Andrei AC, King TE Jr, *et al.* Idiopathic interstitial pneumonia: do community and academic physicians agree on diagnosis? *Am J Respir Crit Care Med* 2007; 175: 1054–1060.
24 Jo HE, Glaspole IN, Levin KC, *et al.* Clinical impact of the interstitial lung disease multidisciplinary service. *Respirology* 2016; 21: 1438–1444.
25 Jo HE, Corte TJ, Moodley Y, *et al.* Evaluating the interstitial lung disease multidisciplinary meeting: a survey of expert centres. *BMC Pulm Med* 2016; 16: 22.
26 Tomassetti S, Wells AU, Costabel U, *et al.* Bronchoscopic lung cryobiopsy increases diagnostic confidence in the multidisciplinary diagnosis of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2016; 193: 745–752.