



Multidisciplinary interobserver agreement in the diagnosis of idiopathic pulmonary fibrosis

M. Thomeer^{*,#}, M. Demedts^{*}, J. Behr[†], R. Buhl⁺, U. Costabel[§], C.D.R. Flower[‡], J. Verschakelen^{*}, F. Laurent^{**}, A.G. Nicholson^{##}, E.K. Verbeken^{*}, F. Capron^{††}, M. Sardina⁺⁺, G. Corvasce⁺⁺ and I. Lankhorst⁺⁺, and the Idiopathic Pulmonary Fibrosis International Group Exploring *N*-Acetylcysteine I Annual (IFIGENIA) study group

ABSTRACT: The purpose of the present study was to evaluate the accuracy of the diagnosis of idiopathic pulmonary fibrosis (IPF) by respiratory physicians in six European countries, and to calculate the interobserver agreement between high-resolution computed tomography reviewers and histology reviewers in IPF diagnosis.

The diagnosis of usual interstitial pneumonia (UIP) was assessed by a local investigator, following the American Thoracic Society/European Respiratory Society consensus statement, and confirmed when a minimum of two out of three expert reviewers from each expert panel agreed with the diagnosis. The level of agreement between readers within each expert panel was calculated by weighted kappa.

The diagnosis of UIP was confirmed by the expert panels in 87.2% of cases. A total of 179 thoracic high-resolution computed tomography scans were independently reviewed, and an interobserver agreement of 0.40 was found. Open or thoracoscopic lung biopsy was performed in 97 patients, 82 of whom could be reviewed by the expert committee. The weighted kappa between histology readers was 0.30.

It is concluded that, although the level of agreement between the readers within each panel was only fair to moderate, the overall accuracy of a clinical diagnosis of idiopathic pulmonary fibrosis in expert centres is good (87.2%).

KEYWORDS: Idiopathic pulmonary fibrosis, kappa coefficient, lung biopsy, radiology

Idiopathic pulmonary fibrosis (IPF) is a specific form of a chronic fibrosing interstitial pneumonia limited to the lung, and is typically characterised by the histological appearance of usual interstitial pneumonia (UIP) on open or thoracoscopic lung biopsy (OLB and TLB, respectively) [1]. The clinical diagnosis of IPF is based on the exclusion of known causes of interstitial lung disease, a restrictive lung function pattern with impaired gas exchange and the presence of a typical pattern of bibasilar reticular abnormalities with minimal ground-glass opacities on thoracic high-resolution computed tomography (HRCT) [1].

Patients with IPF show worse survival than those with other types of idiopathic interstitial pneumonia [2–4]. Since the diagnosis of IPF depends upon the expertise of the pathologist and radiologist, it is important that the clinician knows

the diagnostic accuracy of thoracic HRCT and of lung biopsy in UIP. Various studies have calculated the accuracy of thoracic HRCT in fibrotic lung diseases [5–7], evaluated interobserver agreement for the diagnosis of different thoracic HRCT patterns (*e.g.* ground-glass and reticular pattern) [8, 9] in patients with biopsy-proven nonspecific interstitial pneumonia (NSIP) or UIP [3], or in different forms of interstitial lung disease [10, 11]. Studies on interobserver agreement amongst pathologists are sparse [12], and only one study with a multicentric prospective design has addressed the issue of diagnostic accuracy in UIP in relation to both radiologist and pathologist [7]. No study has addressed this issue in view of the new American Thoracic Society (ATS)/European Respiratory Society (ERS) consensus criteria [1]. Therefore, the aim of the present study was to evaluate the diagnostic accuracy of respiratory physicians in

AFFILIATIONS

^{*}University Hospitals, Catholic University of Leuven, Leuven, and [#]East Limburg Hospital, Genk, Belgium.

[†]Dept of Internal Medicine I, Grosshadern Clinic, Ludwig Maximilian University, Munich,

⁺Dept of Internal Medicine III, Johannes Gutenberg University Clinic, Mainz, and

[§]Dept of Respiratory Medicine, Ruhrland Clinic, Essen, Germany.

[‡]Evelyn Hospital, Cambridge, and

^{##}Royal Brompton Hospital, London, UK.

^{**}Cardiologal Hospital, Bordeaux University Hospital, Bordeaux, and

^{††}Pitié-Salpêtrière Hospital, Paris, France.

⁺⁺Zambon Group, Bresso, Italy.

CORRESPONDENCE

M. Thomeer, UZ Leuven, Afdeling Longziekten, Herestraat 49, B - 3000 Leuven, Belgium. Fax: 32 16346803
E-mail: michiel.thomeer@scarlet.be

Received: May 11 2006

Accepted after revision:

November 21 2007

STATEMENT OF INTEREST

Statements of interest for M. Thomeer, M. Demedts, C.D.R. Flower, J. Verschakelen, F. Laurent, A.G. Nicholson, E.K. Verbeken, F. Capron, M. Sardina, G. Corvasce and I. Lankhorst, and the Idiopathic Pulmonary Fibrosis International Group Exploring *N*-Acetylcysteine I Annual (IFIGENIA) study can be found at www.erj.ersjournals.com/misc/statements.shtml

European Respiratory Journal

Print ISSN 0903-1936

Online ISSN 1399-3003

IPF, and to calculate the interobserver agreement between HRCT reviewers and histology reviewers in the diagnosis of UIP.

METHODS

Patients

All of the patients presented in the current study were included in the Idiopathic Pulmonary Fibrosis International Group Exploring *N*-Acetylcysteine I Annual (IFIGENIA) trial [13]. The IFIGENIA trial is a European prospective double-blind placebo-controlled trial studying the effect of high-dose *N*-acetylcysteine in combination with standard therapy (prednisone and azathioprine) in patients with IPF. Following the judgment of a local investigator, patients were included if the diagnosis of IPF was based on the international consensus criteria [1], and they were aged 18–75 yrs. Newly diagnosed (<6 months) as well as previously diagnosed (>6 months) patients were considered for the study. The IFIGENIA trial was approved by the local ethical committee of the participating centres and every patient signed their informed consent.

HRCT scanning protocol

HRCT of the thorax was performed in the supine position during breath-holding at full inspiration, with 1 or 1.5-mm-thick sections at 1-cm intervals throughout the entire thorax. Images were reconstructed using a high-frequency algorithm at window levels appropriate for the pulmonary parenchyma (mean -500– -700 HU; width 1,400–2,000 HU). No intravenous contrast was administered.

Review by the radiology committee

The local investigator provided copies of the original HRCT scan and sent these to the international trial coordinator (G. Corvasce). A copy was sent to each of the three members of the radiology committee (C.D.R. Flower, F. Laurent and J. Verschakelen). The copies of the HRCT scans were reviewed independently without knowledge of clinical, physiological or pathological parameters. The international trial coordinator ensured that the three members of the radiology committee (reviewers A, B and C) were unaware of the patient's identity. Each member of the committee confirmed the diagnosis of UIP on thoracic HRCT based on the criteria of the international consensus statement [1]. The degree of confidence in the diagnosis was recorded in terms of the scan being very suggestive, probable or unlikely for the diagnosis. The UIP diagnosis on thoracic HRCT was confirmed if the scan was scored as very suggestive or probable for UIP, and rejected if it was scored as unlikely. If disagreement occurred between the three members of the radiology committee, the UIP diagnosis agreed by the majority of the three members was accepted as definite.

Review of lung biopsy specimens by the histology committee

The diagnosis of UIP according to the criteria of the ATS/ERS consensus classification [1] was assessed by an independent panel of three pathology experts (A.G. Nicholson, E.K. Verbeken and F. Capron). The local investigator sent OLB or TLB slides to the international trial coordinator, who blinded the cases and sent them to two members of the pathology review committee (reviewers D and E). All slides were graded as being very suggestive, probable or unlikely for the diagnosis

of UIP. For each observer, the UIP diagnosis on lung biopsy was confirmed if the slide was scored as very suggestive or probable for UIP and rejected if it was scored as unlikely. If the two reviewers disagreed as to diagnosis of UIP, the slides were sent to the third member of the pathology committee (reviewer F) and assessed in an identical fashion. The diagnosis agreed by the majority of the three members was accepted as final. The slides were reviewed independently without knowledge of clinical or physiological parameters.

Definite diagnosis of UIP

The diagnosis of UIP was rejected when one or both committees did not confirm a diagnosis of UIP.

Statistics

Weighted kappa coefficients (κ_w) were used to measure the level of interobserver agreement. The κ_w were calculated using a method recommended for comparing level of agreement with categorical data [14], along with their respective 95% confidence intervals (CIs).

RESULTS

A total of 36 local investigators from six European countries were included (table 1), giving 179 HRCT scans and 82 OLB or TLB specimens for review.

Radiology reviewer A reviewed 178 HRCT scans (one scan was never reviewed), reviewer B 176 (two scans were judged uninterpretable and one was never reviewed) and reviewer C 176 (two scans were judged uninterpretable and one was never reviewed; fig. 1). After combining the observations of all three radiologists, the 532 HRCT observations were judged to be unlikely in 67 (12.6%) cases, probable in 203 (38.2%) and very suggestive in 258 (48.5%) for the diagnosis of UIP. For four (0.8%) observations, the HRCT scan was judged to be uninterpretable because of lack of quality. A total of 238 HRCT observations could be correlated with the results of lung biopsy. When the HRCT scans were judged to be unlikely, probable and very suggestive for UIP, 67.5, 84.4 and 91.7%, respectively, of the corresponding lung biopsy specimens were positive for UIP (fig. 2).

All 82 biopsy specimens (44 OLB and 38 TLB) were sent to the international trial coordinator for review by pathology reviewers D and E. After combining the observations of the three histology reviewers, the 178 OLB/TLB observations were judged to be unlikely in 33 (18.5%) cases, probable in 66 (37.1%) and very suggestive in 76 (42.7%) for the diagnosis of UIP. For three (1.7%) observations, the biopsy slide was judged to be uninterpretable. Reviewer D reviewed all 82 OLB/TLB specimens and reviewer E 79 (three were judged uninterpretable). Histology reviewer F was solicited to review 14 biopsy slides (fig. 1).

In 12.8% of the patients, the diagnosis of UIP was rejected by at least one review committee (table 1; fig. 1). The diagnosis of UIP was confirmed by the pathology review committee for 84% of the 82 OLB/TLB specimens. The diagnosis of UIP on HRCT was confirmed for 92.7% of the 165 HRCT scans (fig. 1).

Table 2 summarises the level of agreement between the three different HRCT reviewers; κ_w ranged 0.33–0.46. No important differences in κ_w were seen within the different subgroups.

TABLE 1 Diagnosis of idiopathic pulmonary fibrosis (IPF) in different subgroups

	IPF by HRCT		UIP by OLB/TLB		Definite diagnosis [#]	
	Absent	Present	Absent	Present	Absent	Present
Subjects n	179		82		179	
Belgium	2 (10.0)	18 (90.0)	1 (8.3)	11 (91.7)	2 (10.0)	18 (90.0)
France	4 (10.0)	36 (90.0)	4 (26.7)	11 (73.3)	6 (15.0)	34 (85.0)
Germany	3 (4.8)	59 (95.2)	5 (29.4)	12 (70.6)	8 (12.9)	54 (87.1)
Italy	1 (5.9)	16 (94.1)	1 (14.3)	6 (85.7)	1 (5.9)	16 (94.1)
The Netherlands	2 (10.5)	17 (89.5)	0 (0.0)	19 (100.0)	2 (10.5)	17 (89.5)
Spain	2 (9.5)	19 (90.5)	3 (25.0)	9 (75.0)	4 (19.0)	17 (81.0)
Total	14 (7.3)	165 (92.7)	14 (16.0)	68 (84.0)	23 (12.8)	156 (87.2)

Data are presented as n (%), unless otherwise indicated. HRCT: high-resolution computed tomography; UIP: usual interstitial pneumonia; OLB: open lung biopsy; TLB: thoracoscopic lung biopsy. [#]: defined as agreement of the histology and radiology committees with the diagnosis of IPF based on OLB/TLB and thoracic HRCT, or of the radiology committee when only HRCT scans were available.

Table 3 summarises the level of agreement between the two pathology reviewers; a κ_w of 0.30 (95% CI 0.12–0.48) was calculated. The level of agreement was 0.84 (0.55–1.14) in the subgroup of those patients in whom the diagnosis of UIP was not confirmed on HRCT. When the severity of lung function impairment (forced vital capacity of >60% or <60% of the predicted value) was taken into account, no difference in level of agreement was observed.

DISCUSSION

Two salient findings emerge from the present study. First, the diagnosis of IPF proposed by a respiratory specialist was rejected in 12.8% of cases after review of histology and HRCT by expert committee. Secondly, the mean level of agreement between the three different HRCT reviewers was 0.40, and between the two pathology reviewers 0.30.

The diagnostic accuracy of a pulmonary physician in IPF in relation to the ATS/ERS diagnostic criteria [1] remains to be established. A confident diagnosis of IPF proposed by a clinician was confirmed in 87.2% of cases in the present study. The rejection of the diagnosis was not based on clinical criteria,

but rather on HRCT and/or lung biopsy findings that were not compatible with the diagnosis of UIP. HUNNINGHAKE *et al.* [7] found the probability of a patient being given a confident diagnosis by the referring clinician to be 81%, similar to the present results. Although the study of HUNNINGHAKE *et al.* [7] represented the first published prospective multicentric study regarding the level of agreement between clinicians, radiologists and pathologists as to the diagnosis of IPF, it is not clear from their study on which clinical grounds the diagnosis of IPF was made, since no clinical or radiological criteria were provided for the diagnosis of IPF.

The present study also addressed the question of agreement between histology reviewers in the diagnosis of UIP in view of the new pathological classification [1]. The interobserver agreement between the histology reviewers was low, with a mean κ_w of 0.30, a level scored as showing fair agreement following the proposed interpretation of κ_w of BRENNAN and

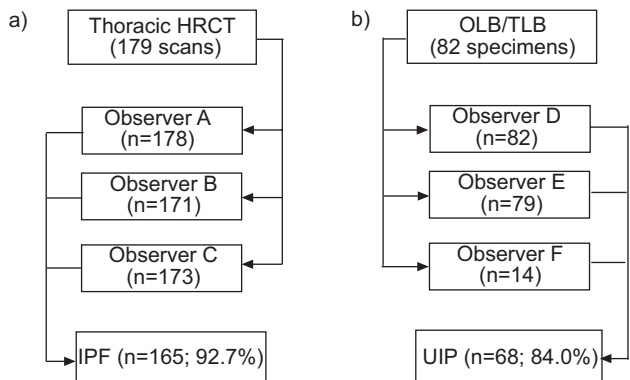


FIGURE 1. Study profile of the a) radiology panel and b) histology panel. Observer F was only solicited when observers D and E disagreed. HRCT: high-resolution computed tomography; IPF: idiopathic pulmonary fibrosis; OLB: open lung biopsy; TLB: thoracoscopic lung biopsy; UIP: usual interstitial pneumonia.

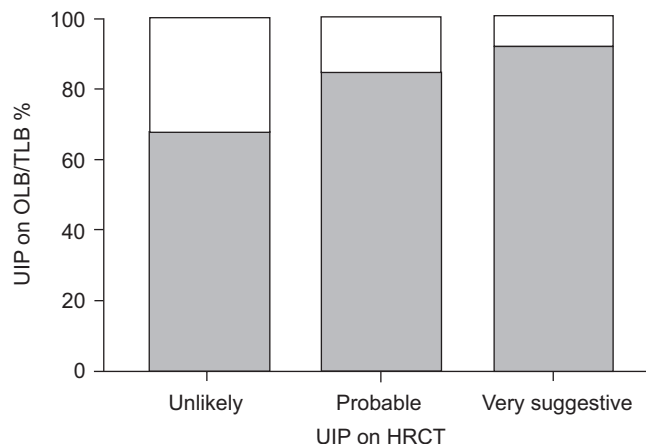


FIGURE 2. The proportion of usual interstitial pneumonia (UIP) diagnosis (□: negative; ■: positive) on open lung biopsy (OLB)/thoracoscopic lung biopsy (TLB) correlated with the degree of confidence of UIP diagnosis on 238 high-resolution computed tomography (HRCT) observations (40 unlikely, 90 probable and 108 very suggestive).

TABLE 2 Weighted kappa coefficients (κ_w) between thoracic high-resolution computed tomography (HRCT) reviewers[#]

	Mean κ_w	κ_w (95% CI)		
		A versus B	A versus C	B versus C
Total	0.40	0.40 (0.29–0.52)	0.33 (0.23–0.44)	0.46 (0.36–0.56)
OLB/TLB: UIP				
Absent	0.41	0.47 (0.03–0.92)	0.24 (-0.21–0.69)	0.52 (0.19–0.86)
Present	0.35	0.35 (0.16–0.54)	0.21 (0.03–0.40)	0.49 (0.33–0.66)
Total reviewed	0.40	0.41 (0.24–0.58)	0.27 (0.11–0.43)	0.53 (0.38–0.67)
TBLB/no biopsy	0.40	0.40 (0.24–0.56)	0.40 (0.26–0.54)	0.39 (0.26–0.53)
FVC				
<60% pred	0.40	0.36 (0.16–0.56)	0.33 (0.15–0.51)	0.50 (0.34–0.67)
>60% pred	0.39	0.42 (0.28–0.56)	0.33 (0.20–0.46)	0.43 (0.30–0.56)

CI: confidence interval; OLB: open lung biopsy; TLB: thoracoscopic lung biopsy; UIP: usual interstitial pneumonia; TBLB: transbronchial lung biopsy; FVC: forced vital capacity; % pred: % predicted. [#]: each reviewer (A, B and C) scored the HRCT scans as unlikely, probable or very suggestive for the diagnosis of UIP.

SILMAN [14]. The κ_w ranges 0–1, with 0 indicating only chance agreement and 1 perfect agreement [14]. NICHOLSON *et al.* [15] and CHERNIACK *et al.* [16] studied levels of agreement between pathologists for individual histological parameters (*e.g.* extent of fibrosis) in biopsy specimens showing UIP, and found κ_w ranging 0.56–0.76 and -0.06–0.30, respectively. However, since their κ_w were calculated for a different purpose, their results are not comparable with those of the present study. NICHOLSON *et al.* [17] presented another study examining the prognostic significance of histological patterns of idiopathic interstitial pneumonia. Slides of 37 lung biopsy specimens with UIP, 28 with NSIP and 13 with desquamative interstitial pneumonia or respiratory bronchiolitis-associated interstitial lung disease were reviewed independently by two pulmonary histopathologists. They found an overall κ_w of 0.49, but the level of agreement in distinguishing between UIP and NSIP was 0.26. In view of patient selection, this latter figure is comparable with the present study. It suggests that distinguishing a UIP or NSIP pattern based on histology is difficult, and perhaps more difficult with knowledge of the lobar histopathological

variability in UIP and NSIP in the same lung [18]. This finding was confirmed in a recent study [12] in which observer agreement on UIP as final histological diagnosis of 0.49 and on NSIP of 0.32 were found.

The level of agreement between the HRCT readers was fair to moderate [14]. This κ_w is comparable with those of MACDONALD *et al.* [9] and FLAHERTY *et al.* [3] (table 4), who compared the interobserver agreement on HRCT for patients with NSIP and UIP. HUNNINGHAKE *et al.* [7] found an interobserver agreement of 0.54. Since the radiological criteria for the diagnosis of UIP on HRCT were not mentioned in this latter study, it is not possible to explain the difference between their κ_w and those of MACDONALD *et al.* [9], FLAHERTY *et al.* [3] and the present study. Other studies [8, 10, 19, 20], however, have found a higher level of agreement, but their study populations and the purposes for which observer variability were calculated differed significantly from those of the present study (table 4). AZIZ *et al.* [21] found an observer agreement of 0.48 on the first-choice diagnosis in a cohort of 131 patients with diffuse parenchymal lung disease; the κ_w in the IPF cohort was 0.50.

It is important to emphasise the fact that radiologists with differing levels of experience and expertise may interpret radiographic images differently. The radiologists in the present study are specialists in thoracic imaging and have extensive expertise in the interpretation of HRCT scans. Each reader was blinded to the clinical parameters, and the reading was performed separately, so that the different readers could not influence each other. The κ_w was used to evaluate observer variability in order to remove that component of agreement attributable to chance. Although this method of statistical analysis permits a more accurate assessment of observer variability than unadjusted data, a κ_w may underestimate a high level of agreement [14]. In the present study, the level of agreement between the different readers was unexpectedly low. Might this be due to the high prevalence of the disease in the study population, or might it be observer variation bias? The interpretation of κ_w depends upon the prevalence of the disease, which was high (0.84) in the present study [14]. The

TABLE 3 Weighted kappa coefficients (κ_w) between histology reviewers D and E[#]

	Mean κ_w (95% CI)
Total	0.30 (0.12–0.48)
Thoracic HRCT: UIP	
Absent	0.84 (0.55–1.14)
Present	0.16 (-0.03–0.36)
FVC	
<60% pred	0.33 (0.05–0.61)
>60% pred	0.28 (0.06–0.50)

CI: confidence interval; HRCT: high-resolution computed tomography; UIP: usual interstitial pneumonia; FVC: forced vital capacity; % pred: % predicted. [#]: each reviewer (D and E) scored the open lung biopsy or thoracoscopic lung biopsy specimens as unlikely, probable or very suggestive for the diagnosis of UIP.

TABLE 4 Overview of studies addressing interobserver agreement on thoracic computed tomography (CT) in various forms of pulmonary fibrosis

First author [Ref.]	Year	Interobserver agreement κ coefficient	Study population	Subjects n	Observers n	Comments
GRENIER [10]	1991	0.64–0.78	Sarcoidosis	53	3	Definition of IPF unclear
			Pulmonary fibrosis	33		
			Histiocytosis X	17		
			Other ILD	37		
WELLS [19]	1993	0.58–0.76	Systemic sclerosis	35	2	Interobserver agreement for grading CT appearance and change in nature and extent of disease
IPF	21					
COLLINS [8]	1994	0.48	Systemic sclerosis	63	4	Interobserver agreement for CT pattern type
IPF	63					
KAZEROONI [20]	1997	0.51–0.83	UIP; DIP	24; 1	4	Interobserver agreement for pattern type in different lobes
MACDONALD [9]	2001	0.40	NSIP	21	4	Interobserver agreement for NSIP and UIP
			UIP	32		
HUNNINGHAKE [7]	2001	0.54	IPF	54	4	Interobserver agreement for IPF versus non-IPF; criteria for IPF diagnosis not mentioned
			Non-IPF	37		
FLAHERTY [3]	2003	0.43	NSIP	23	2	Interobserver agreement for NSIP and UIP
			UIP	73		
AZIZ [21]	2004	0.50	DPLD	131	11	Interobserver agreement for first-choice diagnosis of IPF
Present study		0.40	UIP	156	3	Interobserver agreement for IPF versus non-IPF; IPF patients included following ATS/ERS criteria
			Non-UIP	23		

ILD: interstitial lung disease; IPF: idiopathic pulmonary fibrosis; UIP: usual interstitial pneumonia; DIP: desquamative interstitial pneumonia; NSIP: nonspecific interstitial pneumonia; DPLD: diffuse parenchymal lung disease; ATS: American Thoracic Society; ERS: European Respiratory Society.

prevalence of the disease was high because the HRCT scans and lung biopsy slides were from patients selected by a local investigator who had already confirmed the diagnosis of IPF to conform to the ATS/ERS criteria. The higher prevalence of the disease in the present study population is a possible explanation for the conflicting finding that 67% of histologically confirmed UIP cases gave HRCT results that were reported as unlikely. However, this is not astonishing, since a recent publication reported that 59% of patients with definite or probable NSIP on CT had a histological diagnosis of UIP [4]. The present authors assume that many of the CT scans that had been reported as being unlikely for UIP would fulfil the CT criteria for NSIP.

The results of the present study may evoke concerns about diagnostic accuracy in IPF. This form of lung fibrosis is a rare disease, and no single accurate diagnostic test for it exists. Studies of diagnostic accuracy in IPF are performed mostly in tertiary referral centres. Even in these studies, significant interobserver variability exists. In most of these studies, as in the present one, prior knowledge of the presence of a form of interstitial lung disease exists, which may incite observer bias and therefore influence the results in terms of diagnostic accuracy. The incidence of IPF is low in a general pulmonary practice. Diagnostic accuracy (*i.e.* sensitivity and specificity) also depends upon the prevalence of the disease. A lower prevalence of disease results in a higher number of false positive and false negative diagnoses. If very costly therapeutic options come to market in the future, the only means of ensuring the greatest possible diagnostic accuracy in IPF is to refer patients to centres with expertise in pulmonary histology and thoracic imaging and clinical experience of IPF [22].

In summary, it has been shown that the accuracy of a clinical diagnosis of idiopathic pulmonary fibrosis is 87.2%. Given that idiopathic pulmonary fibrosis has such a poor prognosis [2] in relation to other forms of idiopathic interstitial pneumonia [3], it was concluded that the use of independent high-resolution computed tomography and histology panels to ensure accurate diagnosis of idiopathic pulmonary fibrosis, as performed in the Idiopathic Pulmonary Fibrosis International Group Exploring *N*-Acetylcysteine I Annual study [13], is extremely valuable and helps minimise bias. The present study demonstrated that the use of reviewer panels for radiology and histology in idiopathic pulmonary fibrosis trials is feasible. It is important that the clinician knows that an accurate diagnosis of idiopathic pulmonary fibrosis requires specific expertise that is available in tertiary referral centres with the close collaboration of histopathologists, radiologists and clinicians.

ACKNOWLEDGEMENTS

The members of the Idiopathic Pulmonary Fibrosis International Group Exploring *N*-Acetylcysteine I Annual (IFIGENIA) study group are as follows.

Steering committee: J. Behr (Grosshadern Clinic, Ludwig Maximilian University, Munich, Germany); R. Buhl (Johannes Gutenberg University Clinic, Mainz, Germany); U. Costabel (Ruhland Clinic, Essen, Germany); R. Dekhuijzen (Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands); M. Demedts (Chairman) and M. Thomeer (University Hospitals, Catholic University of Leuven, Leuven, Belgium); H.M. Jansen (Academic Medical Centre, Amsterdam, The Netherlands); W. MacNee (University of Edinburgh Medical School, Edinburgh, UK); and B. Wallaert

(Calmette Hospital, Lille Regional University Hospital, Lille, France).

Country coordinators: P. de Vuyst (Erasmus University Hospital, Brussels, Belgium); B. Wallaert (France); J. Behr (Germany); S. Petruzzelli (Cardiothoracic Dept, Pisa University, Pisa, Italy); J.M.M. van den Bosch (St Antonius Hospital, Nieuwegein, The Netherlands); E. Rodríguez-Becerra (Virgen del Rocío University Hospital, Seville, Spain); W. MacNee (UK).

Radiology review committee: C.D.R. Flower (Evelyn Hospital, Cambridge, UK); J. Verschakelen (University Hospitals, Catholic University of Leuven, Leuven, Belgium); F. Laurent (Cardiological Hospital, Bordeaux University Hospital, Bordeaux, France).

Histology review committee: A.G. Nicholson (Royal Brompton Hospital, London, UK); E.K. Verbeken (University Hospitals, Catholic University of Leuven, Leuven); F. Capron (Pitié-Salpêtrière Hospital, Paris, France).

Local investigators. Belgium: M. Demedts, P. de Vuyst, E. Michiels (East Limburg Hospital, Genk), H. Slabbynck (Middelheim General Hospital, Antwerp), M. Thomeer. France: A. Bourdin and P. Chanez (Arnaud de Villeneuve Hospital, Montpellier), J. Cadranet (Tenon Hospital, Paris), P. Camus (Le Bocage University Hospital, Dijon), P. Delaval (Pontchaillou Hospital, Rennes), N. Just and B. Wallaert (Calmette Hospital, Lille Regional University Hospital, Lille, France), J.F. Muir (Bois Guillaume Hospital, Rouen). Germany: U. Costabel, R. Baumgartner (Grosshadern Clinic, Ludwig Maximilian University, Munich), J. Behr, R. Bonnet and I. Mäder (Bad Berka Central Clinic, Bad Berka), R. Buhl, A.M. Kirsten (Johannes Gutenberg University Clinic, Mainz), R. Loddenkemper (Heckeshorn Lung Clinic, Zehlendorf Clinic, Berlin), A. Meyer (Eppendorf University Hospital, Hamburg), J. Müller-Quernheim (Borstel Research Centre, Medical Clinic, Borstel), H. Steveling (Ruhrland Clinic, Essen, Germany), T. Welte (Magdeburg University Clinic, Magdeburg), H. Worth (Clinic Fürth, Fürth). Italy: G. Anzalone (Prato Hospital, Prato), G.B. Bottino (DIMI, Genoa University, Genoa), G. Bustacchini (S. Maria delle Croci Hospital, Ravenna), M. Dottorini (R. Silvestrini Hospital, Perugia), S. Gasparini (Torrette Hospital, Torrette di Ancona), C. Giuntini (Cardiothoracic Dept, Pisa University, Pisa), A. Rossi (IRCCS S. Matteo General Hospital, Pavia), G. Simon (Azienda Ospedaliera Villa Sofia, Palermo). The Netherlands: F. Beaumont (Bosch Medicentrum, Locatie Grootziekengasthuis, Hertogenbosch), M. Drent (Maastricht University Hospital, Maastricht), H.M. Jansen, J.M.M. van den Bosch, and F.J.J. van den Elshout (Rijnstate Hospital, Arnhem). Spain: J. Ancochea Bermudez (Hospital Universitario de la Princesa, Madrid), L. Callol Sanchez (Hospital Universitario Del Aire, Madrid), J.L. Llorente (Hospital De Cruces, Baracaldo-Bilbao), J.M. Rodríguez-Arias and I. Vigil (Hospital Sant Pau, Barcelona), E. Rodríguez-Becerra (Hospital Universitario Virgen del Rocío, Seville).

Zambon personnel and consultants: A. Ardia (consultant), M. Sardina, G. Corvasce, and I. Lankhorst (consultant).

REFERENCES

- 1 Idiopathic pulmonary fibrosis, diagnosis and treatment. International consensus statement. *Am J Respir Crit Care Med* 2000; 161: 646–664.
- 2 Bjraker JA, Ryu JH, Edwin MK, *et al.* Prognostic significance of histopathologic subsets in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 1998; 157: 199–203.
- 3 Flaherty KR, Thwaite EL, Kazerooni EA, *et al.* Radiological versus histological diagnosis in UIP and NSIP: survival implications. *Thorax* 2003; 58: 143–148.
- 4 Flaherty KR, Toews GB, Travis WD, *et al.* Clinical significance of histological classification of idiopathic interstitial pneumonia. *Eur Respir J* 2002; 19: 275–283.
- 5 Tung KT, Wells AU, Rubens MB, Kirk JM, du Bois RM, Hansell DM. Accuracy of the typical computed tomographic appearances of fibrosing alveolitis. *Thorax* 1993; 48: 334–338.
- 6 Lynch DA, Newell JD, Logan PM, King TE Jr, Müller NL. Can CT distinguish hypersensitivity pneumonitis from idiopathic pulmonary fibrosis? *AJR Am J Roentgenol* 1995; 165: 807–811.
- 7 Hunninghake GW, Zimmerman MB, Schwartz DA, *et al.* Utility of a lung biopsy for the diagnosis of idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2001; 164: 193–196.
- 8 Collins CD, Wells AU, Hansell DM, *et al.* Observer variation in pattern type and extent of disease in fibrosing alveolitis on thin section computed tomography and chest radiography. *Clin Radiol* 1994; 49: 236–240.
- 9 MacDonald SL, Rubens MB, Hansell DM, *et al.* Nonspecific interstitial pneumonia and usual interstitial pneumonia: comparative appearances at and diagnostic accuracy of thin-section CT. *Radiology* 2001; 221: 600–605.
- 10 Grenier P, Valeyre D, Cluzel P, Brauner MW, Lenoir S, Chastang C. Chronic diffuse interstitial lung disease: diagnostic value of chest radiography and high-resolution CT. *Radiology* 1991; 179: 123–132.
- 11 Mathieson JR, Mayo JR, Staples CA, Müller NL. Chronic diffuse infiltrative lung disease: comparison of diagnostic accuracy of CT and chest radiography. *Radiology* 1989; 171: 111–116.
- 12 Nicholson AG, Addis BJ, Bharucha H, *et al.* Inter-observer variation between pathologists in diffuse parenchymal lung disease. *Thorax* 2004; 59: 500–505.
- 13 Demedts M, Behr J, Buhl R, *et al.* High-dose acetylcysteine in idiopathic pulmonary fibrosis. *N Engl J Med* 2005; 353: 2229–2242.
- 14 Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. *BMJ* 1992; 304: 1491–1494.
- 15 Nicholson AG, Fulford LG, Colby TV, du Bois RM, Hansell DM, Wells AU. The relationship between individual histologic features and disease progression in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2002; 166: 173–177.
- 16 Cherniack RM, Colby TV, Flint A, *et al.* Quantitative assessment of lung pathology in idiopathic pulmonary fibrosis. The BAL Cooperative Group Steering Committee. *Am Rev Respir Dis* 1991; 144: 892–900.
- 17 Nicholson AG, Colby TV, du Bois RM, Hansell DM, Wells AU. The prognostic significance of the histologic pattern of interstitial pneumonia in patients presenting with the clinical entity of cryptogenic fibrosing alveolitis. *Am J Respir Crit Care Med* 2000; 162: 2213–2217.

- 18 Flaherty KR, Travis WD, Colby TV, *et al.* Histopathologic variability in usual and nonspecific interstitial pneumonias. *Am J Respir Crit Care Med* 2001; 164: 1722–1727.
- 19 Wells AU, Rubens MB, du Bois RM, Hansell DM. Serial CT in fibrosing alveolitis: prognostic significance of the initial pattern. *AJR Am J Roentgenol* 1993; 161: 1159–1165.
- 20 Kazerooni EA, Martinez FJ, Flint A, *et al.* Thin-section CT obtained at 10-mm increments *versus* limited three-level thin-section CT for idiopathic pulmonary fibrosis: correlation with pathologic scoring. *AJR Am J Roentgenol* 1997; 169: 977–983.
- 21 Aziz ZA, Wells AU, Hansell DM, *et al.* HRCT diagnosis of diffuse parenchymal lung disease: inter-observer variation. *Thorax* 2004; 59: 506–511.
- 22 Thomeer M. External validity of randomised controlled trials in idiopathic pulmonary fibrosis. *Eur Respir J* 2006; 27: 1072.