idiopathic pulmonary alveolar proteinosis have been shown to decrease following whole lung lavage (WLL).[13] We demonstrated for the first time that serum KL-6, SP-A, SP-D, CEA and lactate dehydrogenase levels are elevated and are significantly correlated with disease severity in patients with APAP.[4]

Cytokeratins are the main protein components of the cell cytoskeleton and belong to the family of intermediate filaments. Cytokeratin 19 is a low-molecular weight (40 kD) protein and is expressed in simple epithelia, including bronchial epithelial cells.[14] A fragment of cytokeratin 19, called CYFRA 21-1, can be measured in serum and has been evaluated as a useful tumour marker for non-small cell lung cancer.[14] Moreover, it is also known to be elevated in various types of non-malignant lung disorders,[15] such as bacterial pneumonia, tuberculosis, bronchiectasis and interstitial pneumonia, particularly acute interstitial pneumonia.[16] In addition, increased serum CYFRA 21-1 levels and its decrease after WLL have been reported in pulmonary alveolar proteinosis patients.[17] We have previously reported in a pulmonary alveolar proteinosis patient that high levels of serum CYFRA 21-1 declined following successful GM-CSF inhalation therapy, accompanied by an improvement in arterial blood gas levels and the diffusing capacity of carbon monoxide (DLco).[18] Thus, serum CYFRA 21-1 levels may be related to pulmonary alveolar proteinosis severity.

To assess the value of serum CYFRA 21-1 as a biomarker of APAP reflecting disease severity, we measured serum CYFRA 21-1 levels in APAP patients, and examined the relationship between serum CYFRA 21-1 levels and disease severity using respiratory function tests and arterial blood gas analysis. This is the most comprehensive and largest study to assess the role of serum CYFRA 21-1 as biomarker of APAP.

**Table 1** Patient characteristics at the time of diagnosis of APAP

| | Mean ± SE or number of cases |
|---|---|
| Number of patients | 48 |
| Age (years) | 52.9 ± 1.9 |
| Gender (M/F) | 31/17 |
| Smoking history (NS/ES/CS) | 17/20/11 |
| Dust exposure (Yes/No) | 18/29 |
| MRC (0/1/2/3/4) | 9/18/18/3/0 |
| DSS (I/II/III/IV/V) | 4/20/8/9/6 |
| $PaO_2$ (Torr) | 66.5 ± 2.0 |
| $AaDO_2$ (Torr) | 36.3 ± 2.1 |
| VC, %predicted | 87.5 ± 2.9 |
| DLco, %predicted | 54.9 ± 2.9 |
| CYFRA 21-1 (ng/mL) | 9.03 ± 1.64 |
| KL-6 (U/mL) | 7320 ± 1224 |
| SP-D (ng/mL) | 224 ± 24 |
| CEA (ng/mL) | 10.0 ± 1.3 |
| Anti-GM-CSF autoantibody (ng/mL) | 48.5 ± 4.7 |

Normal serum ranges were CYFRA 21-1 (<2.8 ng/mL), KL-6 (<500 U/mL), SP-D (<110 ng/mL) and CEA (<5.0 ng/mL). The serum CYFRA 21-1 levels of 8 cases out of the total 48 cases were within the normal range. The 8 cases were mild or moderate diseases categorized in DSS 1, 2 and 3, and complaint of mild dyspnoea (MRC 0 and 1).

$AaDO_2$, alveolar-arterial oxygen gradient; APAP, autoimmune pulmonary alveolar proteinosis; CEA, carcinoembryonic antigen; CS, current smoker; DLco, diffusing capacity of carbon monoxide; DSS, disease severity score; ES, ex-smoker; GM-CSF, granulocyte-macrophage colony-stimulating factor; KL-6, Krebs von den Lungen-6; MRC, British Medical Research Council score for shortness of breath upon exertion; NS, non-smoker; $PaO_2$, partial pressure of oxygen; SP-D, surfactant protein-D; VC, vital capacity.

## METHODS

### Subjects

Forty-eight consecutive patients diagnosed with APAP at the Kinki-Chuo Chest Medical Center between 2002 and 2010 were enrolled in the study. The diagnosis of APAP was confirmed based on transbronchial lung biopsy, bronchoalveolar lavage, radiological findings and the presence of anti-GM-CSF autoantibody (Table 1). Serum levels of anti-GM-CSF autoantibody were measured at diagnosis by enzyme-linked immunosorbent assay as previously reported with minor modification,[4,19] and its cut-off level to diagnose APAP was 0.5 μg/mL.

Serum CYFRA 21-1 levels of APAP were compared with that of 68 patients with interstitial lung diseases diagnosed at the Kinki-Chuo Chest Medical Center, including idiopathic pulmonary fibrosis ($n = 25$), non-specific interstitial pneumonia ($n = 6$), collagen vascular disease associated interstitial pneumonia ($n = 15$), chronic hypersensitivity pneumonia ($n = 13$) and sarcoidosis ($n = 9$) (Table S1 in the online supporting information).

WLL was performed in 10 patients, and inhaled GM-CSF was administered to 20 patients (Table S3 in the online supporting information). A positive response to the therapy was defined as improvement of alveolar-arterial oxygen gradient ($AaDO_2$) by 10 Torr.

Comprehensive informed consents for measurements of serum biomarkers and comparison with clinical measures were obtained from all subjects. Additional informed consents were obtained on GM-CSF inhalation therapy. The review board of Kinki-Chuo Chest Medical Center approved this prospective study (Approval Number 73, 99).

### Procedure of WLL

WLL was performed as previously described by Ramirez,[20] with slight modifications. The patient was in supine position, and a left-sided double-lumen tube was placed, after which one-lung ventilation with an inspiratory oxygen fraction of 1.0 was performed. Saline, warmed to body temperature, was delivered by gravity, and the lavage fluid was also drained by gravity after chest-wall percussion. Lavage

was completed when the appearance of the lavage fluid turned from milky to clear, and the total saline delivered usually reached 15–20 L for a single lung. WLL for the other side was performed about 2 weeks after the first WLL. Response to WLL was evaluated 1 month after the first WLL.

### GM-CSF inhalation

Twenty patients were treated with GM-CSF inhalation. Three protocols for GM-CSF inhalation were employed in 15 patients, as described by Tazawa *et al.*[21] One patient from our first pilot study[18] was treated with a daily inhalation of 250 μg of GM-CSF every second week for 24 weeks starting on the first week. Two patients from our second pilot study received 125 μg of GM-CSF inhalation daily during the first 6 weeks, and an additional 125 μg of GM-CSF inhalation daily during the second 6 weeks if the change in $AaDO_2$ was ≥10 Torr, or an additional 250 μg of GM-CSF inhalation daily during the second 6 weeks if the change in $AaDO_2$ was <10 Torr. Twelve patients were from a phase II trial.[21] The remaining five APAP patients were treated with 50 μg of GM-CSF inhalation twice on days 1–8 and no GM-CSF inhalation on days 9–14 for 12 weeks, and 50 μg once on days 1–4 and no GM-CSF inhalation on days 5–14 for the next 12 weeks.

### Evaluation of disease severity

Disease severity of the patients were evaluated at the time of diagnosis, and before and after WLL or GM-CSF inhalation using the following parameters: $AaDO_2$, partial pressure of oxygen ($PaO_2$), percentage of predicted DLco (%DLco), percentage of predicted vital capacity, the British Medical Research Council score for shortness of breath upon exertion and the disease severity score of APAP, as defined by Inoue *et al.*[4] The severity was classified into five grades: grade 1, $PaO_2 \geq 70$ Torr without respiratory symptoms; grade 2, $PaO_2 \geq 70$ Torr with respiratory symptoms; grade 3, 70 Torr > $PaO_2 \geq 60$ Torr; grade 4, 60 Torr > $PaO_2 \geq 50$ Torr; and grade 5, $PaO_2 < 50$ Torr. Arterial blood gas analyses were performed on samples obtained with the patient breathing room air at rest in the supine position for at least 15 min.

### Measurement of CYFRA 21-1 and other serum markers

Serum CYFRA 21-1, KL-6, SP-D and CEA levels were measured by commercial enzyme-linked immunosorbent assay kits (CYFRA, Boehringer Mannheim, Tokyo, Japan; KL-6, Eizai, Tokyo, Japan; SP-D, Kyowa Medex, Tokyo, Japan; CEA, Abbott Japan, Tokyo, Japan). These markers were measured at diagnosis, and before and after therapy.

### Immunohistochemistry

To detect the location of CYFRA 21-1 in the lung, we performed immunohistochemistry using the modified method of Nakayama *et al.*[15] Immunohistochemistry was performed by the avidin-biotin peroxidase complex method using the VECTASTAIN ABC mouse IgG Kit (Vector Laboratories, Burlingame, CA, USA) and a murine monoclonal anti-cytokeratin-19 antibody, Ks 19.1 (Progen, Biotechnik GmbH, Heidelberg, Germany), which is used for the measurement of serum CYFRA 21-1, at the concentration of 5 μg/mL. As a negative control, mouse IgG was used at the same concentration.

### Statistical analysis

Each parameter was presented as mean ± SE. Correlations between serum CYFRA 21-1 and parameters of disease severity were evaluated by Spearman's rank correlation analysis. The differences in serum CYFRA 21-1 levels before and after GM-CSF inhalation or after WLL were determined by the paired *t*-test. Serum CYFRA 21-1 levels of GM-CSF responders and non-responders were compared by the Student *t*-test. Receiver operator characteristic curve analysis was performed to evaluate serum markers as diagnostic tests for APAP. The correlation between the change in serum CYFRA 21-1 and the change in $AaDO_2$ was assessed using the Pearson correlation analysis. A *P*-value of <0.05 was considered statistically significant. All statistical calculations were performed using the JMP version 8.0.2 for Macintosh (SAS Institute Inc., Cary, NC, USA).

## RESULTS

### Cut-off level of CYFRA 21-1 for diagnosis of APAP

Serum CYFRA 21-1 levels were significantly elevated in APAP (9.03 ± 1.64 ng/mL) (Table 1) as compared with interstitial lung diseases (2.96 ± 0.22 ng/mL) (Table S1). Cut-off level of serum CYFRA 21-1 to diagnose APAP was 3.80 ng/mL by receiver operator characteristic analysis (Fig. 1). Area under receiver operator characteristic curve of CYFRA 21-1 was similar to that of KL-6 (Fig. 1, Table S2 in the online supporting information).

### Correlation between serum CYFRA 21-1 levels and other parameters reflecting disease severity

Serum CYFRA 21-1 levels were significantly correlated with British Medical Research Council score, disease severity score, $AaDO_2$ and %DLco, but not with percentage of predicted vital capacity. Serum KL-6 and SP-D were significantly correlated with all five parameters (Table 2).

### Change in serum CYFRA 21-1 and $AaDO_2$ levels 1 month after WLL

In all 10 patients treated with WLL, serum CYFRA levels significantly decreased 1 month after WLL (Fig. 2a). $AaDO_2$ before WLL and 1 month after WLL was evaluated in nine patients, except for one patient whose disease was too severe to evaluate $AaDO_2$ under room air conditions. The decrease in $AaDO_2$ in

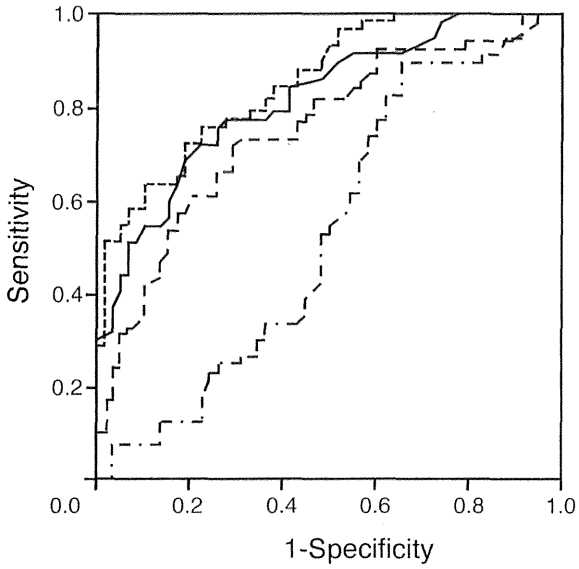


**Figure 1** Receiver operator characteristic (ROC) curves showing comparison of serum markers including KL-6, surfactant protein-D (SP-D), carcinoembryonic antigen (CEA) and CYFRA 21-1 as diagnostic test for autoimmune pulmonary alveolar proteinosis (APAP). Levels of serum markers were measured in 68 patients with interstitial lung diseases (ILD), including idiopathic pulmonary fibrosis; 25 patients, non-specific interstitial pneumonia; 6 patients, collagen vascular disease associated interstitial pneumonia; 15 patients, hypersensitivity pneumonia; 13 patients and sarcoidosis; 9 patients, and 48 patients of APAP. Area under ROC curve of CYFRA 21-1, KL-6, CEA and SP-D was 0.8171, 0.8516, 0.7505 and 0.4951, respectively. Cut-off level of serum CYFRA 21-1 to diagnose APAP was 3.80 ng/mL. Details of serum markers in ILD (Table S1) and results of ROC curve analysis (Table S2) were described in the online supporting information. ——, CYFRA 21-1; ····, KL-6; ·–·–, SP-D; – – –, CEA.

**Table 2** Relationship (ρ) between serum markers and disease severity markers in APAP patients

| | CYFRA 21-1 | KL-6 | SP-D | CEA |
|---|---|---|---|---|
| MRC | 0.6127* | 0.6113* | 0.5117** | 0.5909* |
| DSS | 0.6441* | 0.6991* | 0.5044* | 0.5635* |
| $AaDO_2$ | 0.7474* | 0.7239* | 0.6091* | 0.5914* |
| %DLco | −0.6793* | −0.7309* | −0.6878* | −0.4356*** |
| %VC | −0.2326 | −0.5097** | −0.5222* | −0.2403 |

* $P < 0.0001$; ** $P < 0.001$; *** $P < 0.01$.

The relationship between serum markers and disease severity markers was evaluated by ρ using the Spearman rank correlation analysis.

$AaDO_2$, alveolar-arterial oxygen gradient; APAP, autoimmune pulmonary alveolar proteinosis; CEA, carcinoembryonic antigen; DLco, diffusing capacity of carbon monoxide; DSS, disease severity score; KL-6, Krebs von den Lungen-6; MRC, British Medical Research Council score for shortness of breath upon exertion; SP-D, surfactant protein-D; VC, vital capacity.

nine patients following WLL ($\Delta AaDO_2$) was significantly correlated with the decrease in CYFRA 21-1 after WLL (ΔCYFRA 21-1) (Fig. 2b; $n = 9$, $r = 0.7621$, $P = 0.0170$), but not with ΔKL-6 ($n = 9$, $r = 0.4378$, $P = 0.2386$), ΔSP-D ($n = 9$, $r = 0.4686$, $P = 0.2033$) and ΔCEA ($n = 9$, $r = 0.3948$, $P = 0.2930$).

### Change in serum CYFRA 21-1 and $AaDO_2$ following GM-CSF inhalation therapy

In GM-CSF-effective cases ($n = 11$), serum CYFRA 21-1 levels diminished significantly after GM-CSF inhalation therapy ($P = 0.002$); however, in GM-CSF-ineffective cases ($n = 9$), serum CYFRA 21-1 levels did not change significantly (Fig. 3). Although the characteristics of responders were not different from that of non-responders, except for age (Table S3), the serum CYFRA 21-1 levels in responders were significantly higher compared with non-responders before GM-CSF inhalation therapy (Fig. 4). There was no significant difference in serum KL-6, SP-D and CEA between responders and non-responders (Table S3). Multivariate logistic regression analysis with a stepwise method to predict effectiveness of GM-CSF inhalation revealed that serum CYFRA level (ng/mL) is a significant predictive factor; however, other factors, including the other serum markers, were insignificant (Table 3, Table S4 in the online supporting information).

### Immunohistochemistry for CYFRA 21-1 detection

Immunohistochemical analysis of the transbronchial lung biopsy specimens obtained from one patient was performed. CYFRA 21-1-positivity was detected in the proteinaceous material in the alveolar spaces and in the hyperplastic alveolar epithelial cells (Fig. 5).

## DISCUSSION

Receiver operator characteristic curve analysis revealed that CYFRA 21-1 is a diagnostic marker of APAP. CYFRA 21-1 is also a sensitive serum marker for APAP reflecting disease severity that is comparable to KL-6 and CEA, two markers that we have previously described.[4,22] In the present study, serum CYFRA 21-1 levels were significantly correlated with other disease severity parameters. Serum CYFRA 21-1 levels significantly decreased in APAP patients in whom $AaDO_2$ improved following GM-CSF inhalation therapy and WLL. Thus, we propose that CYFRA 21-1 is an important serum marker for diagnosis and disease severity of APAP. No significant relationship of CYFRA 21-1 with percentage of predicted vital capacity was found. However, no significant correlation might be a natural consequence of the percentage of predicted vital capacity decreasing below normal levels only in severe APAP, as previously reported.[4]

Elevation of serum markers in APAP patients is associated with impaired metabolism through the dysfunction of alveolar macrophages and through augmented permeability of each marker observed in interstitial lung diseases.[23,24] Increased permeability is necessary for KL-6, a large molecule with a molecular weight of more than 1000 kD,[25] to flow into the blood.

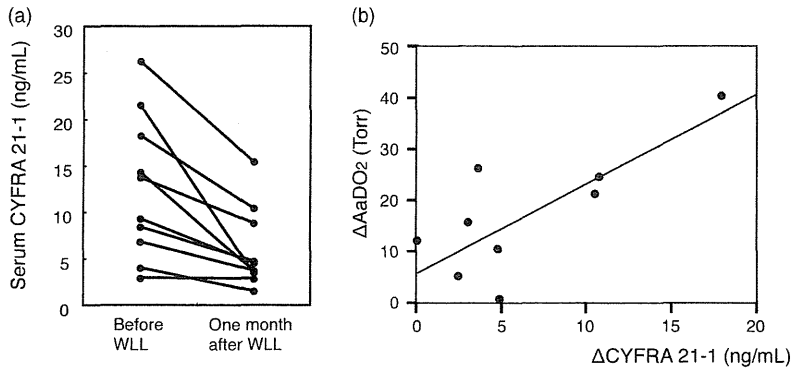


**Figure 2** Serum CYFRA 21-1 levels (mean ± SE) in patients treated with whole lung lavage (WLL) ($n = 10$) decreased significantly after the treatment (from 12.53 ± 2.42 ng/mL to 5.90 ± 1.36 ng/mL) ($P = 0.0041$) (a). Alveolar-arterial oxygen gradient ($AaDO_2$) was evaluated before and after the WLL in nine patients. Decrease in CYFRA 21-1 (ΔCYFRA 21-1) significantly correlated with the decrease in $AaDO_2$ ($\Delta AaDO_2$) in nine patients 1 month after WLL ($P = 0.0170$). Statistical analyses were performed by the paired *t*-test (a) and the Pearson correlation analysis (b).

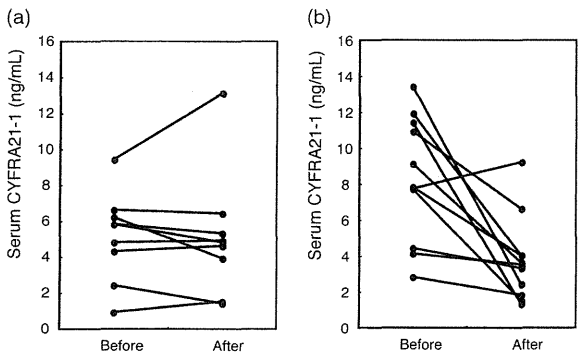


**Figure 3** Measurement of serum CYFRA 21-1 levels (mean ± SE) before and after granulocyte-macrophage colony-stimulating factor (GM-CSF) inhalation. In GM-CSF-ineffective cases (non-responders, $n = 9$), there was no significant difference before (5.13 ± 0.82 ng/mL) and after the inhalation therapy (5.10 ± 1.14 ng/mL) (a), and in GM-CSF-effective cases (responders, $n = 11$), the serum CYFRA 21-1 levels significantly decreased from 8.29 ± 1.04 ng/mL to 3.74 ± 0.71 ng/mL after the inhalation therapy ($P = 0.002$) (b). Statistical analysis was performed by the paired *t*-test.

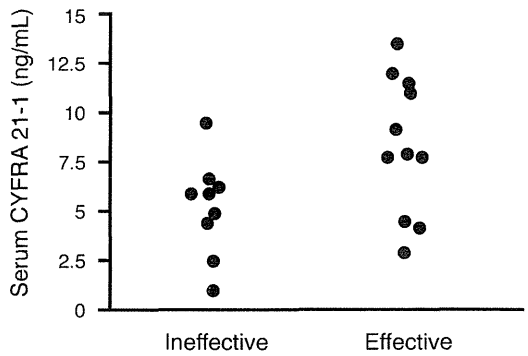


**Figure 4** Serum CYFRA 21-1 levels before the inhalation of granulocyte-macrophage colony-stimulating factor (GM-CSF) were significantly higher in GM-CSF-effective cases (responders, $n = 11$, 8.29 ± 1.04 ng/mL) than those in GM-CSF-ineffective cases (non-responders, $n = 9$, 5.13 ± 0.82 ng/mL), as demonstrated by the Student *t*-test ($P = 0.017$). Three GM-CSF-effective cases showed comparatively low levels of serum CYFRA 21-1, although alveolar-arterial oxygen gradient of three patients was more than 40 Torr. Serum levels of Krebs von den Lungen-6, surfactant protein-D and carcinoembryonic antigen were also low in two out of three patients (data not shown) (values were presented as mean ± SE).

Inoue *et al.* reported that the increase in serum KL-6 level was due to increased permeability of the alveolar wall to blood flow in berylliosis.[26] The permeability of the alveolar-airway barrier might be modulated by GM-CSF. GM-CSF increases alveolar epithelial barrier function in vitro[27] and suppresses the apoptosis of alveolar epithelial cells.[28] Thus, it is possible that the deficiency of GM-CSF caused by the presence of anti-GM-CSF autoantibodies in APAP conversely leads to the dysfunction of the alveolar epithelial barrier and increased permeability.

Based on a previous report on non-malignant pulmonary diseases, elevation of serum CYFRA 21-1 is due to epithelial damage and its increased production in the epithelium.[16] CYFRA 21-1 is expressed in hyperplastic type II pneumocytes and metaplastic cells in patients with idiopathic pulmonary fibrosis.[16] Although the pathophysiology is unknown, alveolar epithelial cell hyperplasia is sometimes observed in the lung specimens of APAP patients.[29] In agreement with this finding, we also pointed out the existence of alveolar epithelial cell hyperplasia in the present study. Thus, increased production of CYFRA 21-1 from alveolar epithelial cells may be another reason for the elevation of serum CYFRA 21-1 levels in APAP patients.

Yoshimasu *et al.* calculated the half-life of CYFRA 21-1 to be 1.5 h from the disappearance curve of serum concentration after the resection of lung cancer.[30] The half-lives of SP-D and KL-6 are not known; however, they are predicted to be longer than that of CYFRA 21-1 when considering their higher molecular weights. A short half-life is an important characteristic as a serum marker because levels of a serum marker with a short half-life can change simultaneously with the change in disease activity. To evaluate this point, the disappearance rate of each serum marker 1 month after WLL was examined in the present study; however, there was no significant difference between the markers (data not shown). This is because 1 month is too long an interval to evaluate the disappearance rate of each of the serum markers.

**Table 3** Multivariate logistic regression analysis to predict effectiveness of GM-CSF inhalation in APAP

| Parameters | Odds ratio | 95% CI | *P*-value |
|---|---|---|---|
| CYFRA 21-1 ≥ 6.4 ng/mL | 9.333 | 1.372–94.188 | 0.0213 |

Multivariate logistic regression analysis with a stepwise method to predict effectiveness of GM-CSF inhalation was performed, using serum levels of CYFRA 21-1, carcinoembryonic antigen (CEA), Krebs von den Lungen-6 (KL-6), surfactant protein-D (SP-D), %vital capacity (VC), % diffusing capacity of carbon monoxide (DLco), alveolar-arterial oxygen gradient ($AaDO_2$), disease severity score (DSS), gender, age and British Medical Research Council score for shortness of breath upon exertion (MRC score). Each parameter was classified into two groups by median except for gender: CYFRA 21-1 ≥ 6.4 ng/mL, CEA ≥ 7.3 ng/mL, KL-6 ≥ 4090 U/mL, SP-D ≥ 213 ng/mL, %VC ≥ 89%, %DLco ≥ 47.5%, $AaDO_2$ ≥ 44 Torr, DSS ≥ 4, MRC ≥ 3, age ≥ 53. Univariate analysis revealed that serum CYFRA 21-1 level was only a significant factor to predict good response of GM-CSF inhalation to GM-CSF inhalation (Table S4 in the online support information). Using stepwise method, serum CYFRA 21-1 level was also only a significant predictor for the response of GM-CSF inhalation therapy and other serum factors were insignificant.

APAP, autoimmune pulmonary alveolar proteinosis; CI, confidence interval; GM-CSF, granulocyte-macrophage colony-stimulating factor.

**Figure 5** Immunohistochemical analysis of transbronchial lung biopsy specimens obtained from one patient. CYFRA 21-1-positivity was observed in lipoproteinaceous material in the alveolar spaces (arrowhead) and in the hyperplastic alveolar epithelial cells (arrow).

Serum CYFRA 21-1 levels in responders to GM-CSF inhalation were significantly higher compared with non-responders. The same tendency has been reported for serum KL-6;[21] however, there was no difference in the levels of the other serum markers between responders and non-responders in our examination. Serum CYFRA 21-1 was a significant factor to predict effectiveness of GM-CSF inhalation therapy by multivariate logistic regression analysis. Hence, serum CYFRA might be a useful serum marker to predict the effectiveness of GM-CSF inhalation. However, studies of additional APAP patients treated with GM-CSF inhalation are needed to make definite conclusions, and response to GM-CSF itself needs to be judged cautiously considering the possibility of spontaneous regression of some cases of APAP.[4]

It remains unresolved from our investigation which is the best serum marker of APAP. KL-6 is a very good marker, and is highly correlated with both symptoms and pulmonary function tests. However, another important ability requested for a good marker is to predict treatment response and prognosis. We expect that serum CYFRA 21-1 might be able to forecast effectiveness of GM-CSF inhalation shown above and disease severity changes, for example spontaneous regression, due to its short half-life. Future studies are needed to reach definite conclusions.

We conclude that serum CYFRA 21-1 is a sensitive and useful serum marker for diagnosis and evaluation of disease severity of APAP. CYFRA 21-1 levels might predict the response to GM-CSF inhalation.

## Acknowledgements

We thank M. Kitaichi and M. Akira for pathological and radiological diagnostic comments; T. Yamasaki, A. Nishiyama, T. Kagawa, A. Hirooka, S. Minomo, Y. Kohashi, M. Tamaya, S. Shiraishi and J. Otsuka for performing the whole lung lavage and GM-CSF inhalation; M. Nakagawa for secretarial work; and N. Sakamoto for data management. All of these persons were in Kinki-Chuo Chest Medical Center. This study was partially supported by grants from the Japanese Ministry of Health, Labour, and Welfare for PAP (H24-Nanchito(Nan)-Ippan-035) (YI, TA, KN); and Grant-in-Aid for Scientific Research, from the Japan Society for the Promotion of Science (Kiban B, Kaigai, B24406027) (YI, KN, MH) and from the National Hospital Organization (YI, TA).

## REFERENCES

1 Rosen SH, Castleman B, Liebow AA. Pulmonary alveolar proteinosis. *N. Engl. J. Med.* 1958; **258**: 1123–42.
2 Kitamura T, Tanaka N, Watanabe J, Uchida J, Kanegasaki S, Yamada Y, Nakata K. Idiopathic pulmonary alveolar proteinosis as an autoimmune disease with neutralizing antibody against granulocyte/macrophage colony-stimulating factor Nakata K. *J. Exp. Med.* 1999; **190**: 875–80.
3 Trapnell BC, Whitsett JA, Nakata K. Pulmonary alveolar proteinosis. *N. Engl. J. Med.* 2003; **349**: 2527–39.
4 Inoue Y, Trapnell BC, Tazawa R, Arai T, Takada T, Hizawa N, Kasahara Y, Tatsumi K, Hojo M, Ichiwata T *et al.* Characteristics of a large cohort of patients with autoimmune pulmonary alveolar proteinosis in Japan. *Am. J. Respir. Crit. Care Med.* 2008; **177**: 752–62.
5 Kuroki Y, Tsutahara S, Shijubo N, Takahashi H, Shiratori M, Hattori A, Abe S, Akino T. Elevated levels of lung surfactant protein A in sera from patients with idiopathic pulmonary fibrosis and pulmonary alveolar proteinosis. *Am. Rev. Respir. Dis.* 1993; **147**: 723–9.
6 Seymour JF, Doyle IR, Nakata K, Presneill JJ, Schoch OD, Hamano E, Uchida K, Fisher R, Dunn AR. Relationship of anti-GM-CSF antibody concentration, surfactant protein A and B

levels, and serum LDH to pulmonary parameters and response to GM-CSF therapy in patients with idiopathic alveolar proteinosis. *Thorax* 2003; **58**: 252–7.
7 Honda Y, Kuroki Y, Matsuura E, Nagae H, Takahashi H, Akino T, Abe S. Pulmonary surfactant protein D in sera and bronchoalveolar lavage fluids. *Am. J. Respir. Crit. Care Med.* 1995; **152**: 1860–6.
8 Hirakata Y, Kobayashi J, Sugama Y, Kitamura S. Elevation of tumor markers in serum and bronchoalveolar lavage fluid in pulmonary alveolar proteinosis. *Eur. Respir. J.* 1995; **8**: 689–96.
9 Takahashi T, Munakata M, Suzuki I, Kawakami Y. Serum and bronchoalveolar fluid KL-6 levels in patients with pulmonary alveolar proteinosis. *Am. J. Respir. Crit. Care Med.* 1998; **158**: 1294–8.
10 Seymour JF, Presneill JJ. Pulmonary alveolar proteinosis. Progress in the first 44 years. *Am. J. Respir. Crit. Care Med.* 2002; **166**: 215–35.
11 Usui Y, Takayama S, Nakayama M, Miura H, Kimula Y. Interstitial lattice shadow and mediastinal lymphadenopathy with an elevation of carcinoembryonic antigen in severe pulmonary alveolar proteinosis. *Intern. Med.* 1992; **31**: 422–5.
12 Nakajima M, Manabe T, Niki Y, Matsushima T. Serum KL-6 level as a monitoring marker in a patient with pulmonary alveolar proteinosis. *Thorax* 1998; **53**: 809–11.
13 Sugimoto C, Arai T, Nishiyama A, Inoue Y, Kagawa T, Akira M, Matsumuro A, Hirose M, Kitaichi M, Hayashi S *et al.* [Multidisciplinary assessment of effects, safety and procedure of whole lung lavage for 8 patients with autoimmune pulmonary alveolar proteinosis]. *Nihon Kokyuki Gakkai Zasshi* 2011; **49**: 569–76.
14 Rapellino M, Niklinski J, Pecchio F, Furman M, Baldi S, Chyczewski L, Ruffini E, Chyczewska E. CYFRA 21-1 as a tumor marker for bronchogenic carcinoma. *Eur. Respir. J.* 1995; **8**: 407–10.
15 Nakayama M, Satoh H, Ishikawa H, Fujiwara M, Kamma H, Ohtsuka M, Sekizawa K. Cytokeratin 19 fragment in patients with nonmalignant respiratory diseases. *Chest* 2003; **123**: 2001–6.
16 Dobashi N, Fujita J, Ohtsuki Y, Yamadori I, Yoshinouchi T, Kamei T, Takahashi J. Elevated serum and BAL cytokeratin 19 fragment in pulmonary fibrosis and acute interstitial pneumonia. *Eur. Respir. J.* 1999; **14**: 574–8.
17 Minakata Y, Kida Y, Nakanishi H, Nishimoto T, Yukawa S. Change in cytokeratin 19 fragment level according to the severity of pulmonary alveolar proteinosis. *Intern. Med.* 2001; **40**: 1024–7.
18 Arai T, Hamano E, Inoue Y, Ryushi T, Nukiwa T, Sakatani M, Nakata K. Serum neutralizing capacity of GM-CSF reflects disease severity in a patient with pulmonary alveolar proteinosis successfully treated with inhaled GM-CSF. *Respir. Med.* 2004; **98**: 1227–30.
19 Kitamura T, Uchida K, Tanaka N, Tsuchiya T, Watanabe J, Yamada Y, Hanaoka K, Seymour JF, Schoch OD, Doyle I *et al.* Serological diagnosis of idiopathic pulmonary alveolar proteinosis. *Am. J. Respir. Crit. Care Med.* 2000; **162**: 658–62.
20 Ramirez RJ. Bronchopulmonary lavage: new techniques and observations. *Dis. Chest* 1996; **50**: 581–8.
21 Tazawa R, Trapnell BC, Inoue Y, Arai T, Takada T, Nasuhara Y, Hizawa N, Kasahara Y, Tatsumi K, Hojo M *et al.* Inhaled granulocyte/macrophage-colony stimulating factor as therapy for pulmonary alveolar proteinosis. *Am. J. Respir. Crit. Care Med.* 2010; **181**: 1345–54.
22 Inoue Y, Nakata K, Arai T, Tazawa R, Hamano E, Nukiwa T, Kudo K, Keicho N, Hizawa N, Yamaguchi E *et al.* Epidemiologic and clinical features of idiopathic pulmonary alveolar proteinosis in Japan. *Respirology* 2006; **11**: S55–S60.
23 Doyle IR, Nicholas TE, Bersten AD. Partitioning lung and plasma proteins: circulating surfactant proteins as biomarker of alveolocapillary permeability. *Clin. Exp. Pharmacol. Physiol.* 1999; **26**: 185–97.
24 Hermans C, Bernard A. A lung epithelium-specific proteins: characteristics and potential applications as markers. *Am. J. Respir. Crit. Care Med.* 1999; **159**: 646–78.
25 Kohno N, Kyoizumi S, Awaya Y, Fukuhara H, Yamakido M, Akiyama M. New serum indicator of interstitial pneumonitis activity. Sialylated carbohydrate antigen KL-6. *Chest* 1989; **96**: 68–73.
26 Inoue Y, Barker E, Daniloff E, Kohno N, Hiwada K, Newman LS. Pulmonary epithelial cell injury and alveolar-capillary permeability in berylliosis. *Am. J. Respir. Crit. Care Med.* 1997; **156**: 109–15.
27 Pelaez A, Bechara RI, Joshin PC, Brown LA, Guidot DM. Granulocyte/macrophage colony-stimulating factor treatment improves alveolar epithelial barrier function in alcoholic rat lung. *Am. J. Physiol. Lung. Cell Mol. Physiol.* 2004; **286**: L106–11.
28 Paine R 3rd, Wilcoxen SE, Morris SB, Sartori C, Baleeiro CO, Matthay MA, Christensen PJ. Transgenic overexpression of Granulocyte Macrophage-colony stimulating factor in the lung prevents hyperoxic lung injury. *Am. J. Pathol.* 2003; **163**: 2397–406.
29 Katzenstein AL. Miscellaneous specific disease of uncertain etiology. In: Katzenstein AL (ed.) *Katzenstein and Askin's Surgical Pathology of Non-Neoplastic Lung Disease*, 3rd edn. W.B. Saunders, Philadelphia, PA, 1997; 393–416.
30 Yoshimasu T, Maebeya S, Suzuma T, Bessho T, Tanino H, Arimoto J, Sakurai T, Naito Y. Disappearance curves for tumor markers after resection of intrathoracic malignancies. *Int. J. Biol. Markers* 1999; **14**: 99–105.

## Supporting information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

**Table S1** Patient characteristics of interstitial lung diseases

**Table S2** Details of ROC curve analysis of serum markers

**Table S3** Patient characteristics before therapy

**Table S4** Univariate logistic regression analysis to predict effectiveness of GM-CSF inhalation in APAP

SARCOIDOSIS VASCULITIS AND DIFFUSE LUNG DISEASES 2014; 31; 19-27 

# THE WASOG SARCOIDOSIS ORGAN ASSESSMENT INSTRUMENT: AN UPDATE OF A PREVIOUS CLINICAL TOOL

*M.A. Judson[1], U. Costabel[2], M. Drent[3], A. Wells[4], L. Maier[5], L. Koth[6], H. Shigemitsu[7], D.A. Culver[8], J.Gelfand[6], D. Valeyre[9], N. Sweiss[10], E. Crouser[11], A. S. Morgenthau[12], E.E. Lower[13], A. Azuma[14], M. Ishihara[15], S. Morimoto[16], T. Yamaguchi[17], N. Shijubo[18], J.C. Grutters[19], M. Rosenbach[20], H.P. Li[21], P. Rottoli[22], Y. Inoue[23], A. Prasse[24], R.P. Baughman[13], The WASOG Sarcoidosis Organ Assessment Instrument Investigators**

[1]Albany Medical College; Albany, New York USA; [2]Ruhrlandklinik, University Hospital, University Duisburg-Essen, Essen, Germany; [3]Gelderse Vallei Hospital Ede; Maastricht University, Maastricht, The Netherlands; [4]Royal Brompton Hospital, London; [5]National Jewish Health, Denver, Colorado USA; [6]University of California, San Francisco, San Francisco, CA USA; [7]University of Southern California Keck School of Medicine, Los Angeles, California USA; [8]Cleveland Clinic, Cleveland, Ohio USA; [9]AP-HP, Hôpital Avicenne, Université Paris 13, Sorbonne Paris Cité, France; [10]University of Illinois College of Medicine at Chicago, Chicago, Illinois USA; [11]The Ohio State University Medical Center, Columbus, Ohio USA; [12]Icahn School of Medicine at Mount Sinai; New York, New York USA; [13]University of Cincinnati Medical Center; Cincinnati, OH, USA; [14]Nippon Medical School, Tokyo, Japan; [15]Department of Ophthalmology and Visual Science, Yokohama City, Japan; [16]Aoyama General Hospital, Tokyo, Japan; [17]Japan Railway Tokyo General Hospital, Tokyo, Japan; [18]JR Sapporo Hospital, Sapporo, Japan; [19]St. Antonius Hospital, Utrecht, The Netherlands; [20]University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania USA; [21]Tongji University, School of Medicine, Shanghi, China; [22]Azienda Ospedaliera Universitaria Senese, Italy; [23]Kinki-Chuo Chest Medical Center, Osaka, Japan; [24]University Hospital, Freiburg, Germany; * see appendix 1

ABSTRACT. *Introduction:* A Case Control Etiology of Sarcoidosis Study (ACCESS) sarcoidosis organ assessment instrument has been used for more than a decade to establish uniform standards for the probability of sarcoidosis organ involvement. The ACCESS instrument has become increasingly outdated as new technologies have been developed. Furthermore, the ACCESS instrument failed to address all possible organs involved with sarcoidosis. For these reasons, the World Association of Sarcoidosis and Other Granulomatous Diseases (WASOG) developed a new sarcoidosis organ assessment instrument. *Methods:* Clinical sarcoidosis experts assessed various clinical manifestations for the probability of sarcoidosis organ involvement. Two criteria were required to apply this assessment: 1) histologic evidence of granulomatous inflammation of unknown cause in an organ that was not being assessed; 2) the clinical manifestation being addressed required that alternative causes other than sarcoidosis had been reasonably excluded. Clinical manifestations were assessed as either: a) highly probable: likelihood of sarcoidosis causing this manifestation of at least 90%.; b) probable: likelihood of sarcoidosis causing this manifestation of between 50 and 90%; c) possible: likelihood of sarcoidosis causing this manifestation of less than 50%. The sarcoidosis experts voted on the likelihood of sarcoidosis causing each manifestation using Delphi study methodology where at least 70% agreement of the experts was needed for consensus. *Results:* Various clinical manifestations were classified as highly probable, at least probable, possible, or indeterminate when no consensus could be reached. *Conclusion:* An instrument was developed by expert opinion that may be useful for the clinician and researcher in establishing criteria for sarcoidosis organ involvement.

*(Sarcoidosis Vasc Diffuse Lung Dis 2014; 31: 19-27)*

Received: 01 November 2013
Accepted after revision: 9 December 2013
Correspondence: Marc A. Judson, M.D.
Division of Pulmonary and Critical Care Medicine MC-91, Albany Medical College, 47 New Scotland Avenue, Albany, New York 12208 USA.
E-mail: judsonm@mail.amc.edu

KEY WORDS: sarcoidosis, organ, diagnosis, consensus

## Introduction

Sarcoidosis is a multisystem granulomatous disease of unknown cause.[1] Granulomatous inflammation from sarcoidosis may occur in any organ. ACCESS (A Case Control Etiology of Sarcoidosis Study) was a study of sarcoidosis in the 1990's funded by the National Institutes of Health that was primarily aimed at searching for the etiology of sarcoidosis. For the purposes of that effort, the ACCESS investigators were concerned with establishing criteria for sarcoidosis organ involvement. They had noticed that a large number of sarcoidosis clinical trials did not describe rigorous entry criteria for enrollment; other trials that did clearly describe entry criteria for various sarcoidosis organ involvements were not consistent with each other. This issue was the major impetus for the development of the ACCESS Sarcoidosis Organ Assessment Instrument.[2] The instrument assessed various clinical findings of 15 organs in terms of their likelihood of representing sarcoidosis. These clinical findings were graded as "definite," "probable," and "possible" evidence of organ involvement with sarcoidosis. For the purposes of the ACCESS study, "definite" and "probable" involvement was considered to represent organ involvement with sarcoidosis. A prerequisite for using the ACCESS instrument to evaluate the likelihood of a clinical finding of an organ representing sarcoidosis was that at least one additional organ had demonstrated granulomatous inflammation of no alternative cause.

The ACCESS instrument, although useful, suffers from some deficiencies. First, it was developed more than one decade ago, and it is somewhat outdated as new technologies have been developed in the interim for the diagnosis and monitoring of sarcoidosis. Second, the instrument failed to cover all possible organs involved with sarcoidosis. Third, several common and very specific manifestations of sarcoidosis were not addressed. For these reasons, the World Association of Sarcoidosis and Other Granulomatous Disorders (WASOG) developed a new sarcoidosis organ assessment instrument. This manuscript will describe this instrument.

## Methods

Clinical sarcoidosis experts who were members of WASOG were invited to participate in development of the instrument (Appendix 1). Individuals were invited to serve as "organ group leaders," who would be responsible for developing criteria for a specific organ. Other WASOG members were invited to participate as members of any organ group that was of interest to them. In addition, the organ group leaders were given the authority to invite non-WASOG members to their groups if they believed such individuals had clinical expertise in the assessment of sarcoidosis involvement of specific organs. A list of the organ group leaders and members is shown in Appendix 1.

Various clinical manifestations were assessed for the probability of organ involvement. Two criteria were required to be fulfilled in order to apply this assessment: 1) histologic evidence of granulomatous inflammation of unknown cause needed to be demonstrated in at least one organ that was not being assessed; 2) the clinical manifestation being assessed required that all alternative causes other than sarcoidosis for this clinical manifestation had been reasonably excluded. Provided that these two criteria were fulfilled, clinical manifestations were assessed by assigning them to one of three categories, highly probable, probable, or possible, interpreted as follows. **Highly probable**: such a manifestation is highly specific for sarcoidosis, with a likelihood of sarcoidosis causing this manifestation of at least 90%. In such cases, organ involvement may be assumed without a biopsy. **Probable**: such a manifestation is fairly specific for sarcoidosis, with a likelihood of sarcoidosis causing this manifestation of between 50 and 89%. In general, organ involvement in this category would be adequate to establish a clinical diagnosis of sarcoidosis in that organ. **Possible**: such a manifestation is not specific for sarcoidosis. For all organs, a biopsy showing granulomatous inflammation where alternative causes were reliably excluded was considered "highly probable"; therefore, such a biopsy was not included in this instrument.

Each organ group developed a list of common clinical conditions in their specific organ that could be considered as representing organ involvement with sarcoidosis. The probability of each of these clinical conditions as representing sarcoidosis was determined through organ group discussion that occurred via email and/or conference calls. Although organ group participants were encouraged to use the limited medical evidence available, these criteria were established by expert opinion. Other groups relied on voting of the members. After all organ assessments were developed

by each organ group, each of the clinical conditions was presented to all the organ group leaders and other sarcoidosis experts (all the authors of this manuscript) for a vote as to whether that clinical condition represented "highly probable," "probable," or "possible" involvement of sarcoidosis in that organ (vide supra). The voting group was blinded as to the assessments made by each specific organ group. It was this final vote by which the determination of likelihood of organ involvement on the basis of each clinical condition was based.

In terms of the voting, Delphi study methodology was used to determine that consensus was reached in that at least 70% of the experts needed to agree for consensus.[3] If at least 70% of the experts voted that the clinical condition was "highly probable" to represent sarcoidosis, a consensus was reached. If fewer than 70% of the experts voted that the clinical condition was "highly probable" but at least 70% of the experts voted that it was "highly probable" or "probable," then a consensus was reached that the clinical condition was at least probable. If at least 70% of the experts voted that the clinical condition was only "possible" to represent sarcoidosis, a consensus was reached. Finally, if less than 70% of the experts agreed that the clinical condition was a) "highly probable," b)"highly probable" or "probable" or c)"possible," then a consensus was not reached. In this case, it was unclear, in the opinion of these experts, if such a clinical condition was adequate or inadequate to represent organ involvement with sarcoidosis. After all these expert opinions were rendered, this manuscript was written and was presented to the Executive Committee of WASOG for editing, comments, and approval.

The following organs were evaluated in this instrument: lung, skin, eye, liver, calcium dysregulation, neural tissue, kidney, heart, peripheral lymph node, bone marrow, spleen, bone/joint, ear-nose-throat, parotid/salivary glands, and muscle. In addition, a category of "other organs" was created to encompass all organs involved other than the 15 specific ones listed above.

## Results

Table 1 shows the characteristic of the experts. All but one cared for more than 50 sarcoidosis patients yearly. Almost all had participated in at least one clinical sarcoidosis trial and more than three-quarters had published more than 10 manuscripts concerning sarcoidosis.

Table 2 shows the voting results and consensuses reached in terms of the likelihood of various clinical manifestations representing specific organ involvement with sarcoidosis. For all of the 16 organs, the experts reached consensus that at least one clinical condition was as "at least 'probable'" as representing sarcoidosis. Although several of these clinical conditions concerned laboratory testing, several of them concerned physical examination findings or patient symptoms.

## Discussion

We have proposed the WASOG Organ Assessment Instrument as an update of the original ACCESS organ assessment instrument[2] based on improvement in diagnostic testing for sarcoidosis and new medical evidence that has occurred over the previous decade. Although this document is based as much as possible on medical evidence, it also incorporates expert opinion of WASOG members and

Table 1. Characteristics of the Experts

| | |
|---|---|
| Number of years since training completed (N, %): | |
| <5 | 2, 8% |
| 5-10 | 3, 12% |
| 10-20 | 6, 23% |
| >20 | 15, 58% |
| Number of sarcoidosis patients treated per year, on average: | |
| 0 – 50 | 1, 4% |
| 51 – 100 | 6, 23% |
| 101 – 250 | 11, 42% |
| >250 | 8, 31% |
| Number of sarcoidosis clinical trials that you have participated in: | |
| 0 | 3, 12% |
| 1-3 | 9, 35% |
| 4-10 | 10, 38% |
| >10 | 4, 15% |
| Number of publications authored or co-authored concerning sarcoidosis: | |
| 0 | 0, 0% |
| 1-10 | 6, 23% |
| 11-20 | 4, 15% |
| >20 | 16, 62% |

Table 2. Histological types of cancer in the "sarcoidosis and cancer" group.

| | Highly Probable | At Least Probable | Possible | No Consensus |
|---|---|---|---|---|
| Lung | CXR: bilateral hilar adenopathy (19-2-0)<br>Chest CT: perilymphatic nodules (18-2-1)<br>Chest CT: symmetrical hilar/mediastinal adenopthy (21-0-0)<br>PET/Gallium-67: mediastinal/hilar enhancement (17-4-0) | CXR: diffuse infiltrates (4-13-3)<br>CXR: upper lobe fibrosis (9-10-2)<br>Chest CT: peribronchial thickening (10-8-3)<br>BAL: lymphocytic alveolitis (6-14-1)<br>BAL: elevated CD4/CD8 ratio (11-9-1)<br>PET/Gallium-67: diffuse parenchymal lung enhancement (9-8-4)<br>TBNA: lymphoid aggregates/giant cells (7-8-5) | CXR: localized infiltrate (1-2-18)<br>PFT: obstruction (1-2-17) | PFT: restriction (2-6-12)<br>PFT: isolated reduction in diffusing capacity (2-6-12) |
| Skin | Lupus pernio (16-2-0) | Subcutaneous nodules or plaques (3-14-1)<br>Inflammatory papules within a scar or tattoo (7-8-2)<br>Violaceous or erythematous annular lesions (2-15-3)<br>Violaceous or erythematous macular, papular lesions around the eyes, nose, or mouth (11-3-3) | Atypical lesions: ulcerative, erythrodermic, alopecic, ichthyosiform (0-2-14) | Verrucous/scaly papules or plaques (3-7-7)<br>Hypo- or Hyperpigmented macules or patches (2-8-7) |
| Liver | | Abdominal imaging demonstrating hepatomegaly (1-12-5)<br>Abdominal imaging demonstrating hepatic nodules (3-13-1) | | Hepatomegaly on physical examination (0-10-8)<br>Serum alkaline Phosphate > 3X the upper limit of normal (3-8-6) |
| Eye | uveitis (16-1-0)<br>optic neuritis (13-2-2)<br>mutton fat keratic precipitates (12-1-3)<br>iris nodules (13-3-0)<br>snowball/string of pearls (pars planitis) (10-3-1) | lacrimal gland swelling (10-4-3)<br>trabecular meshwork nodules (9-6-0)<br>retinitis (5-9-0)<br>scleritis (5-7-2)<br>multiple chorioretinal peripheral lesions (6-8-1)<br>adnexal nodularity (8-5-2)<br>candle wax drippings (11-2-3) | cataract (0-1-15)<br>glaucoma (1-3-12)<br>red eye (0-3-14) | blindness (0-7-10)<br>painful eye (0-5-10)<br>cystoid macular edema (2-7-5) |
| Spleen | | Low attenuation nodules on CT (7-11-1)<br>PET/gallium-67 uptake in splenic nodules (7-11-1)<br>Splenomegaly on imaging or physical examination (4-11-4) | | |
| Salivary Gland | Positive gallium-67 scan ("Panda sign") (14-3-0)<br>Positive PET scan of the parotid glands (12-5-0) | Symmetrical parotitis with syndrome of mumps (10-5-2)<br>Enlarged salivary glands (2-10-5) | Dry mouth (2-3-12) | |

*Continued*

Table 2. (Continued)

| | | | | |
|---|---|---|---|---|
| ENT | | Granulomatous changes on direct laryngoscopy (12-5-1)<br>Consistent imaging studies (e.g. sinonasal erosion, mucoperiosteal thickening, positive PET scan) (6-8-4) | chronic sinusitis (0-1-17) | Nasal crusting, epistaxis, or anosmia associated with chronic sinus congestion (1-9-8) |
| Calcium-VitD | hypercalcemia plus all of the following: a) a normal serum PTH level; b) a normal or increased 1,25-OH dyhydroxy vitamin D level; c) a low 25-OH vitamin D level (17-2-0)<br>hypercalciuria plus all of the following: a) a normal serum PTH level; b) a normal or increased 1,25-OH dyhydroxy vitamin D level; c) a low 25-OH vitamin D level (16-3-0) | nephrolithiasis plus all of the following: a) a normal serum PTH level; b) a normal or increased 1,25-diOH vitamin D level; c) a low 25-OH vitamin D level (12-7-0)<br>hypercalciuria without serum PTH and 25 and 1,25 vitamin D levels (3-11-5)<br>nephrolithiasis with calcium stones, without serum PTH and 25 and 1,25 vitamin D levels (4-11-4) | Nephrolithiasis, no stone anlaysis (1-3-15) | |
| Bone-Joint | Typical radiographic features (trabecular pattern, osteolysis, cysts/punched out lesions) (16-3-0) | Dactylitis (10-6-2)<br>Nodular tenosynovitis (4-9-4)<br>Positive PET, MRI, or gallium-67 bone imaging (8-9-1) | Arthralgias (0-5-14) | Non-specific arthritis (1-6-12) |
| Bone Marrow | PET displaying diffuse uptake (13-4-2) | | | leukopenia (2-8-8)<br>anemia (1-5-13)<br>thrombocytopenia (1-5-13) |
| Muscle | | Positive imaging (MRI, Gallium-67) (13-7-0)<br>Palpable muscle masses (3-11-6) | Myalgias (0-6-14) | Elevated serum muscle enzymes (5-8-7) |
| Extra-Thoracic Lymph Node | | Multiple enlarged palpable cervical or epitrochelar lymph nodes without B symptoms (5-13-1)<br>Enlarged lymph nodes identified by imaging in at least 2 peripheral or visceral lymph node stations without B symptoms (5-14-1) | | Multiple enlarged palpable peripheral or visceral lymph nodes with B symptoms (1-10-9)<br>Multiple palpable enlarged peripheral or visceral lymph nodes at sites other than cervical and epitrochlear (2-10-7) |
| Kidney | | Treatment-responsive renal failure with no other risk factors. (9-9-1)<br>Treatment-responsive renal failure in patient with diabetes and/or hypertension. (0-12-5) | Renal failure with other potential risk factors (0-4-15) | CT evidence of abnormal renal enhancement. (0-12-7) |

*Continued*

Table 2. (Continued)

| | | | | |
|---|---|---|---|---|
| Nervous System | Clinical syndrome consistent with granulomatous inflammation of the meninges, brain, ventricular (CSF) system, cranial nerves, pituitary gland, spinal cord, cerebral vasculature or nerve roots<br>-plus-<br>An abnormal MRI characteristic of neurosarcoidosis, defined as exhibiting abnormal enhancement following the administration of gadolinium or a cerebrospinal fluid exam demonstrating inflammation (17-3-0) | Isolated facial palsy, negative MRI (6-8-5)<br>Clinical syndrome consistent with granulomatous inflammation of the meninges, brain, ventricular (CSF) system, cranial nerves, pituitary gland, spinal cord, cerebral vasculature, nerve roots but without characteristic MRI or CSF findings (4-11-4) | Seizures, negative MRI (0-3-15)<br>Cognitive decline, negative MRI (0-1-17) | Peripheral neuropathy involving large fibers (including axonal and demyelinating polyneuropathies and multiple mononeuropathies) (4-9-6)<br>Cranial nerve palsies other than VII, negative MRI (4-7-8)<br>Pleocytosis in the CSF (1-7-10)<br>Low CSF glucose (0-6-12) |
| Cardiac | | Treatment responsive CM or AVNB (12-7-1)<br>Reduced LVEF in the absence of other clinical risk factors (2-13-4)<br>Spontaneous or inducible sustained VT with no other risk factor (6-12-1)<br>Mobitz type II or 3rd degree heart block (11-6-2)<br>Patchy uptake on dedicated cardiac PET (10-8-1)<br>Delayed enhancement on CMR (12-5-1)<br>Positive gallium uptake (8-11-0)<br>Defect on perfusion scintigraphy or SPECT scan (4-11-3)<br>T2 prolongation on CMR (2-11-5) | Reduced LVEF in the presence of other risk factors (e.g., HTN, DM) (0-1-17)<br>Atrial dysrhythmias (0-4-15) | Frequent ectopy (>5% QRS) (0-6-13)<br>Bundle branch block (2-8-9)<br>Impaired RV function with a normal PVR (0-8-10)<br>Fragmented QRS or pathologic Q waves in ≥2 anatomically contiguous leads (0-7-10)<br>At least one abnormal SAECG domain (0-6-10)<br>Interstitial fibrosis or monocyte infiltration (4-8-7) |
| Other Organs | | Positive imaging (3-8-3) | | |

*: *at least 70% agreement by the experts*

**: *for all clinical conditions, a) biopsy of that organ demonstrating granulomatous inflammation of no alternate cause implies highly probable involvement, b) another organ has demonstrated granulomatous inflammation of no alternate cause, c) alternative causes for the clinical manifestation have been reasonable excluded; CXR: chest radiograph; PFT: pulmonary function tests; Chest CT: chest computed tomography scan; TBNA: transbronchial needle aspiration (of a mediastinal lymph node); PET: positron emission tomography scan; Gallium-67: Gallium-67 nuclear scan; BAL: bronchoalveolar lavage; 3X: three times; PTH: serum parathyroid hormone; ENT: ear, nose, throat; Vit D: vitamin D; OH: hydroxy; di-OH: di-hydroxy; MRI: magnetic resonance imaging; B symptoms: fever, weight loss, or night sweats; CSF: cerebral spinal fluid; CM: cardiomyopathy; AVNB: atrioventricular nodal block; LVEF: left ventricular ejection fraction; HTN: systemic hypertension; DM: diabetes mellitus; VT: ventricular tachycardia; RV: right ventricular; SAECG: signal-averaged electrocardiogram; CMR: cardiovascular magnetic resonance imaging*

others who have expertise in the various organ manifestations of sarcoidosis. Unlike the original ACCESS organ assessment instrument, an organized process including a blinded vote was used to reach a consensus of the experts. In addition, the category of "definite" organ involvement in the ACCESS instrument was changed to "highly probable" because even histologic evidence of granulomatous inflammation is not definitive for the diagnosis of sarcoidosis.

This instrument should be viewed as a tool to assign probability to specific clinical findings as representing organ involvement with sarcoidosis. Many believe that because sarcoidosis is a multisystem disease, the diagnosis requires the presence of granulomatous inflammation in at least two organs.[1,4] It is unclear if this requirement is universally agreed upon. Regardless, this instrument assigns a probability for an additional organ having sarcoidosis based on clin-

ical criteria if another organ has demonstrated granulomatous inflammation of unknown cause previously. For clinicians who require that two organs demonstrate granulomatous inflammation of unknown cause for sarcoidosis to be diagnosed, this instrument would allow the diagnosis of sarcoidosis to be established in many cases without the need to biopsy a second organ.

This instrument is not designed to be used to assess activity or severity of sarcoidosis. Furthermore, this instrument is not a suggested algorithm to detect specific sarcoidosis organ involvement. In most cases, sarcoidosis organ involvement that does not cause significant symptoms does not require therapy.[5] Therefore, there is little reason in most cases to pursue a diagnosis of sarcoidosis in every possible organ that may be involved with the disease. Organ involvement may be occult without causing any clinical manifestations, and we are not advocating using this instrument to determine if organ involvement is clinically significant. In addition, the instrument is not designed to determine the need for treatment. It may be appropriate to treat clinical findings meeting only possible involvement criteria if the clinician determines that this is warranted.

This instrument may give guidance as to whether a clinical diagnosis of sarcoidosis organ involvement can be made without performing a biopsy to demonstrate granulomatous inflammation. Taking these individual clinical scenarios in isolation without regard to other clinical findings, we would propose that highly probable or at least probable involvement suggests that scenario is adequate for a clinical diagnosis of organ involvement. We acknowledge that the presence of a scenario voted as possible involvement may be adequate for a clinical diagnosis of sarcoidosis if additional other clinical findings are present.

There are several limitations of this instrument. First, the likelihood of each clinical finding described in the instrument is assigned a probability of representing sarcoidosis involvement of an organ based on the assumption that all other alternative causes for that clinical finding have been "reasonably excluded." This instrument provides no metric for this process, so that the method of excluding alternative diagnoses is arbitrary. At a minimum, attempts should be made to exclude mycobacterial infection, fungal infection, and malignancy. We acknowledge that if a very rigorous process is made to exclude alternative causes for the clinical findings discussed, that the likelihood of sarcoidosis could potentially be "upgraded." Second, there is no evidence that this instrument identifies sarcoidosis phenotypes that relate to specific genotypes or other specific mechanisms of disease. Other instruments have demonstrated evidence of such associations, albeit weakly.[6] It is possible that this instrument might function similarly, but that remains conjectural at this time. Third, the organ manifestations of sarcoidosis in our instrument are not comprehensive. Therefore, several manifestations were not appraised by the experts and, therefore, are unclassified. In addition, this instrument did not evaluate "para-sarcoidosis syndromes" that are often of major concern to sarcoidosis patients. These are conditions found frequently in sarcoidosis patients but are not directly attributable to granulomatous organ involvement and include small fiber neuropathy,[7-9] fatigue,[10-13] depression[13-15] and constitutional symptoms such as fever, weight loss, and malaise.[16] Finally, similar to our comments concerning the need for a biopsy in the preceding paragraph, each of the clinical manifestations that we assessed in this instrument does not always occur in isolation. It is possible that if a patient has evidence of multiple manifestations, each of which we regard as "probable" sarcoidosis, this may raise the probability of sarcoidosis to "highly probable." However, such an analysis is too complex to be examined presently.

We acknowledge that our position that highly probable or probable organ involvement is adequate for a clinical diagnosis of sarcoidosis involvement in an organ is arbitrary. Some may prefer to be more rigorous and require that organ involvement be highly probable for sarcoidosis organ involvement to be assumed without performing a confirmatory biopsy. For these reasons, Table 1 supplies the votes of all the experts for each clinical condition and designates the clinical conditions where a consensus was reached that they were highly probable.

In summary, we have presented an instrument that we consider useful in assessing the probability of organ involvement with sarcoidosis. Although we believe that this instrument will be useful for the clinician and clinical researcher involved with sarcoidosis patients, we suspect that it will require further modification over time as additional diagnostic tests are developed and new medical evidence is generated.

## Appendix 1: Organ Groups for initial establishment of clinical scenarios for future voting

**Lung**
Lead: Robert Baughman. Members: Norman Soskel, Athol Wells, Elliott Crouser, Laura Koth, Marjolein Drent, Paola Rittoli, Daniel Culver, Milton Rossman, Ulrich Costabel, Lisa Maier, Dominique Valeyre, Hide Shigemitsu, Nadera Sweiss, Dominique Israel-Biet, Manuel Riberto Neto, Dheeraj Gupta, Eva Carmona; Patterson, Karen, Andrew P. Matragrano

**Skin**
Lead: Misha Rosenbach. Members: Marc Judson, Gloria Westney, Debasis Sahoo

**Eye**
Lead: Robert Baughman. Members: Elyse Lower, Adam Morgenthau

**Liver**
Lead: Adam Morgenthau. Members: Marjolein Drent, Gloria Westney, Lisa Maier; Nadera Sweiss

**Calcium**
Lead: Marc Judson. Members: Lisa Maier, Laura Koth, Hide Shigemitsu; Nadera Sweiss

**Neuro**
Lead: Jeffery Gelfand. Members: Marjolein Drent, Barney Stern, Jinny Tavee, Elske Hoitsma, Hide Shigemitsu, Kenkichi Nozaki, Fleur Cohen Aubart

**Kidney**
Lead: Elliott Crouser. Members: Milton Rossman, Daniel Culver; Nadera Sweiss

**Heart**
Lead: Daniel Culver. Members: Elliott Crouser, Nabeel Hamzeh, Milton Rossman, Ulrich Costabel, Vasanth Vedantham, Lisa Maier, Adam Morgenthau, Catherine Chapelon, David Bernie, Debabrata Bandyopadhyay

**Peripheral Lymph Node**
Lead: Lower. Member: Marc Judson

**Bone Marrow**
Lead: Adam Morgenthau. Members: Elyse Lower

**Spleen**
Lead: Elyse Lower. Members: Gloria Westney, Adam Morgenthau

**Bone/Joint**
Lead: Nadera Sweiss. Members: Laura Koth, Debasis Sahoo, Andrew Gross ; Arthur Yee

**ENT**
Lead: Marc Judson. Members: Gloria Westney, Lisa Maier

**Parotid/Salivary Glands**
Lead: Robert Baughman. Member: Debasis Sahoo

**Muscle**
Lead: Dominique Valeyre. Members: Marjolein Drent, Nadera Sweiss, Arthur Yee

**Other organs**
Lead: Marc Judson. Member: Robert Baughman

## References


1. Hunninghake GW, Costabel U, Ando M, et al. ATS/ERS/WASOG statement on sarcoidosis. American Thoracic Society/European Respiratory Society/World Association of Sarcoidosis and other Granulomatous Disorders. Sarcoidosis Vasc Diffuse Lung Dis 1999; 16:149-173
2. Judson MA, Baughman RP, Teirstein AS, et al. Defining organ involvement in sarcoidosis: the ACCESS proposed instrument. ACCESS Research Group. A Case Control Etiologic Study of Sarcoidosis. Sarcoidosis Vasc Diffuse Lung Dis 1999; 16:75-86
3. de Meyrick J. The Delphi method and health research. Health Educ 2003; 103:7-16
4. Judson MA. The diagnosis of sarcoidosis. Clin Chest Med 2008; 29:415-427, viii
5. Baughman RP, Culver DA, Judson MA. A concise review of pulmonary sarcoidosis. Am J Respir Crit Care Med 2011; 183:573-581
6. Prasse A, Katic C, Germann M, et al. Phenotyping sarcoidosis from a pulmonary perspective. Am J Respir Crit Care Med 2008; 177:330-336

7. Hoitsma E, Drent M, Verstraete E, et al. Abnormal warm and cold sensation thresholds suggestive of small-fibre neuropathy in sarcoidosis. Clin Neurophysiol 2003; 114:2326-2333
8. Bakkers M, Merkies IS, Lauria G, et al. Intraepidermal nerve fiber density and its application in sarcoidosis. Neurology 2009; 73:1142-1148
9. Judson MA. Small fiber neuropathy in sarcoidosis: Something beneath the surface. Respir Med 2011; 105:1-2
10. Drent M, Lower EE, De Vries J. Sarcoidosis-associated fatigue. Eur Respir J 2012; 40:255-263
11. de Kleijn WP, De Vries J, Lower EE, et al. Fatigue in sarcoidosis: a systematic review. Curr Opin Pulm Med 2009; 15:499-506
12. Wirnsberger RM, de Vries J, Wouters EF, et al. Clinical presentation of sarcoidosis in The Netherlands an epidemiological study. Neth J Med 1998; 53:53-60
13. de Kleijn WP, Drent M, De Vries J. Nature of fatigue moderates depressive symptoms and anxiety in sarcoidosis. Br J Health Psychol 2013; 18:439-452
14. Chang B, Steimel J, Moller DR, et al. Depression in sarcoidosis. Am J Respir Crit Care Med 2001; 163:329-334
15. Elfferich MD, De Vries J, Drent M. Type D or 'distressed' personality in sarcoidosis and idiopathic pulmonary fibrosis. Sarcoidosis Vasc Diffuse Lung Dis 2011; 28:65-71
16. Demirkok SS, Basaranoglu M, Akinci ED, et al. Analysis of 275 patients with sarcoidosis over a 38 year period; a single-institution experience. Respir Med 2007; 101:1147-1154

## Erratum corrige

In the issue 4-2013 of Sarcoidosis Vasculitis and Diffuse lung Diseases in the article "Role of Propionibacterium acnes in sarcoisosis: a meta-analysis "by Y. Zhou, Y. Hu, H. Li, The correct Corresponding Autor is: Huiping Li, MD E.mail: liw2013@126.com

Respiratory Medicine (2014) 108, 1023–1030

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/rmed

# Design of the INPULSIS™ trials: Two phase 3 trials of nintedanib in patients with idiopathic pulmonary fibrosis

Luca Richeldi [a,*], Vincent Cottin [b], Kevin R. Flaherty [c], Martin Kolb [d], Yoshikazu Inoue [e], Ganesh Raghu [f], Hiroyuki Taniguchi [g], David M. Hansell [h], Andrew G. Nicholson [h], Florence Le Maulf [i], Susanne Stowasser [j], Harold R. Collard [k]

[a] *National Institute for Health Research Southampton Respiratory Biomedical Research Unit and University of Southampton, University Road, Southampton SO17 1BJ, UK*
[b] *Louis Pradel Hospital, University of Lyon, 28 Avenue du Doyen Lepine, 69677 Bron Cedex, Lyon, France*
[c] *University of Michigan Health System, 1500 E. Medical Center Drive, 3916 Taubman Center, Ann Arbor, MI 48109-0360, USA*
[d] *McMaster University, Department of Medicine, Pathology & Molecular Medicine, 50 Charlton Avenue East, Hamilton, Ontario L8N 4A6, Canada*
[e] *National Hospital Organization Kinki-Chuo Chest Medical Center, Department of Diffuse Lung Diseases and Respiratory Failure, Clinical Research Center, 1180 Nagasone-cho, Kita-Ku, Sakai City, Osaka 591-8555, Japan*
[f] *University of Washington, Seattle, WA 98105, USA*
[g] *Tosei General Hospital, Department of Respiratory Medicine and Allergy, 160 Nishioiwake-cho, Seto, Aichi 489-8642, Japan*
[h] *Royal Brompton and Harefield Hospital NHS Foundation Trust and National Heart and Lung Institute, Imperial College, Sydney Street, London SW3 6NP, UK*
[i] *Boehringer Ingelheim France S.A.S., 12, rue André Huet – B.P. 292, 51060 Reims Cedex, France*
[j] *Boehringer Ingelheim Pharma GmbH & Co. KG, Binger Str. 173, 55216 Ingelheim, Germany*
[k] *University of California San Francisco, 505 Parnassus Avenue, San Francisco, CA 94131, USA*

Received 1 March 2014; accepted 16 April 2014
Available online 29 April 2014

**KEYWORDS**
Clinical trial;
Phase III;
Drug therapy;

**Summary**
*Background:* Nintedanib is in clinical development as a treatment for idiopathic pulmonary fibrosis (IPF). Data from the Phase II TOMORROW study suggested that nintedanib 150 mg twice daily had clinical benefits with an acceptable safety profile.

* Corresponding author. Tel.: +44 (0) 23 8120 6663; fax: +44 (0) 23 8051 1761.
*E-mail address:* L.Richeldi@soton.ac.uk (L. Richeldi).

http://dx.doi.org/10.1016/j.rmed.2014.04.011

Protein kinase inhibitor;
Protein tyrosine kinases

*Methods:* The INPULSIS™ trials are replicate Phase III, randomized, double-blind, studies comparing the efficacy and safety of nintedanib 150 mg twice daily with placebo in patients with IPF. Eligible patients were aged ≥40 years with a diagnosis of IPF within 5 years before randomization who had undergone a chest high-resolution computed tomography (HRCT) scan within 1-year before screening, and who had a forced vital capacity (FVC) of ≥50% predicted and a diffusing capacity for carbon monoxide of 30–79% predicted. Participants were randomized 3:2 to receive nintedanib or placebo for 52 weeks. The primary endpoint is the annual rate of decline in FVC. The key secondary endpoints are change from baseline in the total score on the St. George's Respiratory Questionnaire (a measure of health-related quality of life) over 52 weeks and time to first acute exacerbation.

*Results:* Enrolment of 1066 patients in 24 countries was completed in September 2012. Results will be reported in the first half of 2014.

*Conclusion:* The INPULSIS™ trials will determine the efficacy of nintedanib in patients with IPF, including its impact on disease progression as defined by decline in FVC, acute exacerbations and health-related quality of life. In addition, they will characterise the adverse event profile of nintedanib in this patient population.

**Trial registration**: Registered at ClinicalTrials.gov (identifiers: NCT01335464 and NCT01335477).



## Introduction

Idiopathic pulmonary fibrosis (IPF) is a specific form of chronic, progressive fibrosing interstitial pneumonia [1]. An accurate diagnosis of IPF requires the exclusion of other known causes of interstitial lung disease, the presence of a specific radiological pattern of usual interstitial pneumonia (UIP) determined by high-resolution computed tomography (HRCT), or specific combinations of HRCT and histopathologic patterns in patients who have undergone surgical lung biopsy [1]. IPF is considered a rare disease [2]. In a retrospective cohort study conducted in the United States using data from a large healthcare claims database spanning a 5-year period, the prevalence of IPF was estimated to be 14 to 43 cases per 100,000, and the annual incidence to be 6.8 to 16.3 per 100,000, depending on how cases were defined [3]. Similarly, in the United Kingdom, the annual incidence of IPF was estimated to be 7.4 per 100,000 based on primary care data from 2000 to 2008 [4]. IPF is ultimately a fatal disease, with a reported median survival time of approximately 3 years from diagnosis [5]. In addition, the symptoms of IPF impact negatively on patients' physical function and emotional well-being, as well as their health-related quality of life (HRQoL) [6,7].

An improved understanding of the pathogenic mechanisms underlying IPF over the last decade has resulted in several agents being evaluated in clinical trials [8] and in pirfenidone being approved for the treatment of a subgroup of patients with IPF in several countries. Results of four large randomized, double-blind, placebo-controlled Phase III trials investigating the efficacy and safety of treatments for IPF are awaited this year: the PANTHER-IPF trial of N-acetylcysteine (NAC) (NCT00650091), the ASCEND trial of pirfenidone (NCT01366209), and the INPULSIS™ trials of nintedanib (NCT01335464 and NCT01335477).

Nintedanib (formerly known as BIBF 1120) is a potent tyrosine kinase inhibitor targeting intracellular receptors of fibroblast growth factor receptor (FGFR), platelet-derived growth factor receptor (PDGFR), and vascular endothelial growth factor receptor (VEGFR) [9]. Activation of these receptor kinases has been implicated in multiple pathways in the pathogenesis of IPF [10,11]. *In vitro* studies and animal models suggest that nintedanib has anti-fibrotic and anti-inflammatory effects that may attenuate the progression of fibrosis [12,13]. Results from the Phase II TOMORROW trial suggested that 12 months' treatment with nintedanib 150 mg twice daily results in a reduced rate of decline in forced vital capacity (FVC), fewer acute exacerbations and preservation of HRQoL, measured using the St. George's Respiratory Questionnaire (SGRQ) [14]. The purpose of this manuscript is to describe the design of the INPULSIS™ studies, two replicate Phase III trials that further investigate the efficacy and safety of nintedanib 150 mg twice daily compared with placebo in patients with IPF.

## Methods

### Trial design

Both the INPULSIS™ trials are multinational, randomized, double-blind, parallel-group studies comparing the efficacy and safety of nintedanib 150 mg twice daily with placebo in patients with IPF. The INPULSIS™ trials were initiated in May 2011 and enrolment ($n$ = 1066) was completed in September 2012. Patients were recruited in 24 countries in the Americas, Europe, Asia and Australia. Following a screening period, eligible patients were randomized 3:2 (using an interactive phone/web response system) to receive nintedanib or placebo for 52 weeks (Fig. 1). Each study concluded with a 4-week follow-up period after completion of the 52-week treatment period. A 3:2 ratio was chosen to aid enrolment. In order to reduce the amount of missing data, patients who discontinued trial drug, for any reason, prior to completing the 52 weeks' treatment were asked to attend all visits and undergo all examinations

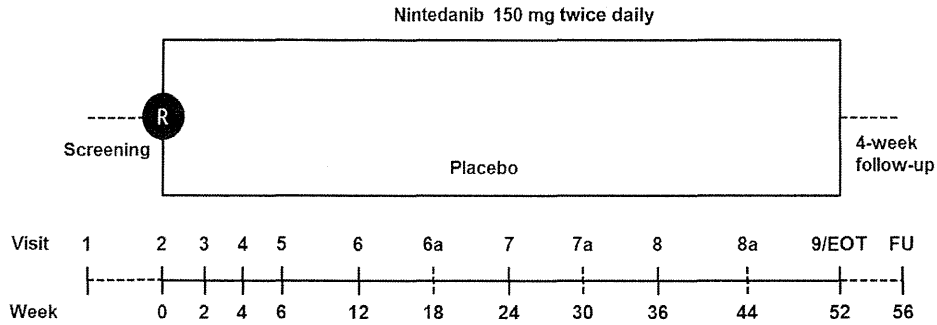


**Figure 1** INPULSIS™ trial design. R, randomization (3:2 ratio for nintedanib:placebo); EOT, end of treatment; FU, follow-up. FVC was measured at all visits except visits 6a, 7a and 8a, which were for blood sampling for laboratory tests only.

as originally planned. In addition, vital status at week 52 was to be collected for all patients who prematurely discontinued but did not agree to attend all visits until week 52.

For each trial, the sample size was calculated to provide 90% power to detect a difference of 100 mL/year between the treatment groups in the rate of FVC decline. Based on the Phase II TOMORROW trial data, the common standard deviation for change from baseline in FVC was assumed to be 300 mL. Assuming data from 2% of patients would be non-evaluable, the sample size was calculated as 194 patients in the placebo group and 291 patients in the nintedanib 150 mg twice daily group if using a 2 group *t*-test at a 1-sided 2.5% level. Since the primary analysis is a random coefficient regression model, including adjustment for several variables and taking into account information across time rather than at a single time-point, it is expected that the power will be greater than the 90% calculated for the *t*-test.

As in the Phase II TOMORROW trial, dose interruption and/or reduction of the dose from 150 mg twice daily to 100 mg twice daily was allowed for the management of adverse events. After an adverse event had resolved, the dose could be reinstituted at 150 mg twice daily. The investigators were provided with guidelines on the management of diarrhoea, a known side-effect related to treatment with tyrosine kinase inhibitors [15,16]. Guidelines on the management of liver enzyme elevations were also provided to the investigators. Patients who completed the 52-week treatment period and the 4-week follow-up period in the INPULSIS™ trials were invited to participate in an open-label extension trial (NCT01619085).

## Trial organisation and oversight

The INPULSIS™ trials were guided by an advisory committee consisting of clinical experts in IPF and representatives of the sponsor, Boehringer Ingelheim. An independent Data Monitoring Committee (DMC) regularly reviewed the data, in particular serious adverse events, adverse events leading to discontinuation of study drug, and laboratory parameters, and made recommendations to the sponsor about the continuation of the trials. An Adjudication Committee reviewed medical documentation for all deaths to evaluate the primary cause of death in a blinded manner. This committee also adjudicated all events reported by the investigators as meeting the criteria for an acute exacerbation of IPF as defined in the protocol, classifying them as a confirmed acute exacerbation, suspected acute exacerbation, or not an acute exacerbation.

Both trials were conducted in accordance with the principles of the Declaration of Helsinki and the Harmonized Tripartite Guideline for Good Clinical Practice from the International Conference on Harmonization and were approved by local authorities. The clinical trial protocol was approved by an Independent Ethics Committee and/or Institutional Review Board at all the participating centres. All patients provided written informed consent prior to study entry.

## Patients

To be eligible to participate in the INPULSIS™ trials, patients had to be ≥40 years of age with a diagnosis of IPF established within 5 years before randomization, to have undergone chest HRCT within 12 months before screening, and to have an FVC ≥50% of predicted value [17] and a carbon monoxide diffusion capacity ($DL_{CO}$) of 30–79% of predicted value [18]. The diagnosis of IPF was established based on the central review of chest HRCT scans from all patients by an expert radiologist (DMH) according to protocol-specified criteria (Table 1). Surgical lung biopsy specimens were also centrally evaluated if available by an expert pathologist (AGN).

**Table 1** Diagnostic criteria for IPF based on chest HRCT if surgical lung biopsy was not available. To qualify for a diagnosis of IPF if a surgical lung biopsy was not available, the criteria A and B and C; or criteria A and C; or criteria B and C had to be met.

| | |
|---|---|
| A | Definite honeycomb lung destruction with basal and peripheral predominance |
| B | Presence of reticular abnormality and traction bronchiectasis consistent with fibrosis with basal and peripheral predominance |
| C | Atypical features are absent, specifically nodules and consolidation. Ground glass opacity, if present, is less extensive than reticular opacity pattern |

Patients with abnormal laboratory parameters (liver transaminases or bilirubin above 1.5-fold upper limit of normal), cardiac disease (i.e. myocardial infarction within 6 months or unstable angina within 1 month of randomization), or who, in the opinion of the investigator, were likely to receive a lung transplant during the study were not permitted to enter the trial. Patients who were taking full-dose anticoagulant therapy or high-dose antiplatelet therapy at screening, or had received treatment with NAC or prednisone >15 mg/day or equivalent within 2 weeks of screening, or pirfenidone, azathioprine, cyclophosphamide, cyclosporine A or any investigational drug within 8 weeks of screening, were excluded. Concomitant therapy with prednisone ≤15 mg/day or equivalent was permitted if the dose had been stable for ≥8 weeks prior to screening. Patients who experienced deterioration, as judged by the investigator, were permitted to receive concomitant treatment with azathioprine, cyclophosphamide, cyclosporine A, NAC, or prednisone >15 mg/day or equivalent at the discretion of the investigator 6 months or more after starting to receive study medication. In cases of acute exacerbation, any treatments could be freely initiated or increased as deemed appropriate by the investigator. However, pirfenidone and any investigational treatments for IPF were not allowed throughout the trial.

## Outcome measures

The primary endpoint for the INPULSIS™ trials is the annual rate of decline in FVC (mL/year), calculated from measurements obtained over the 52 weeks of treatment (Fig. 2). Spirometry testing was conducted according to ATS/ERS criteria, including daily calibration of the spirometer, regular calibration of the calibration pump and FVC tests conducted in triplicate, with the highest result selected [19]. All spirometry was performed on sponsor-provided machines and ongoing feedback and training were provided.

The key secondary endpoints are change from baseline in SGRQ total score over 52 weeks and time to first acute exacerbation. Acute exacerbations were defined as events meeting all of the following criteria: unexplained worsening or development of dyspnoea within 30 days, new diffuse pulmonary infiltrates on chest X-ray and/or HRCT, or parenchymal abnormalities with no pneumothorax or pleural effusion (new ground-glass opacities) since last visit. Causes of the acute worsening, including infection, left heart failure, pulmonary embolism or any identifiable cause of acute lung injury were to be excluded as per routine clinical practice and microbiological studies. Investigator-reported exacerbations were adjudicated by the Adjudication Committee. Other secondary endpoints include absolute changes from baseline in FVC (mL and % predicted); proportion of FVC responders (patients who did not have an absolute decline in FVC % predicted of >5% or >10%); risk of an acute exacerbation; change from baseline in $SpO_2$ (oxygen saturation) at rest; change from baseline in $DL_{CO}$ at rest (measured in accordance with ATS/ERS guidelines [20]); all-cause, respiratory, and 'on-treatment' time to death. Composite endpoints of time to death or lung transplant, and time to death or lung transplant or meeting arbitrary pre-defined criteria for lung transplant (FVC <45% predicted *or* $DL_{CO}$ <30% predicted *or* $SpO_2$ <88% at rest) were also included in order to capture a range of outcomes indicating an unfavourable clinical course.

Further patient-reported outcomes (PROs) investigated in the INPULSIS™ trials are the change from baseline to week 52 in the score on the three SGRQ domains (impact, symptoms, activity) [21], SGRQ-I [22], University of California San Diego Shortness of Breath Questionnaire [UCSD-SOBQ] [23], EuroQol 5-dimensional quality of life questionnaire [EQ5D], Cough and Sputum Assessment Questionnaire cough domains [CASA-Q(CD)] [24]; the proportion of 4-point responders on SGRQ total score; and the proportion of responders on Patient's Global Impression of Change (PGI-C). Safety assessment will include reporting of

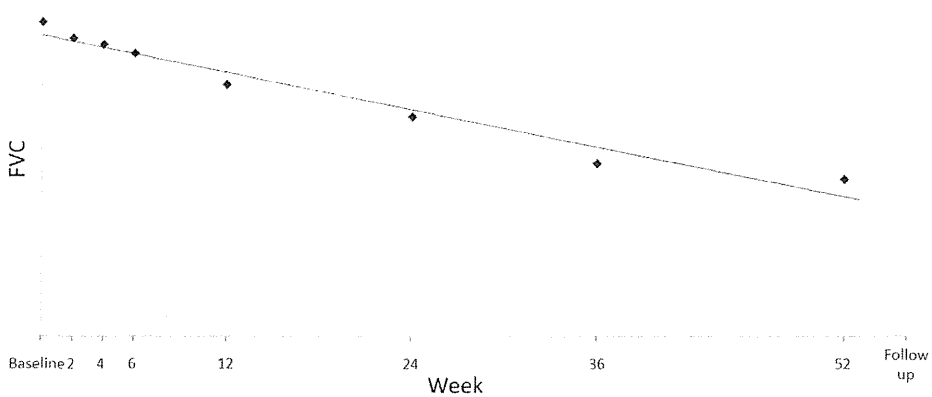


**Figure 2** Methodology for calculating slope of FVC decline. The primary endpoint will be analysed using a random coefficient regression (random slopes and intercepts) model, including gender, age and height as covariates. Visits are planned at 2, 4, 6, 12, 24, 36 and 52 weeks after randomization. All available FVC values except the value from the follow-up visit will be used in this analysis except for patients who prematurely discontinue trial medication, in which case the value from the follow-up visit will also be used.

adverse events; assessment of vital signs, physical examination and weight; clinical laboratory tests (haematology, clinical chemistry and urinalysis).

### Statistical analysis

Efficacy and safety analyses will be conducted on patients who were randomized to treatment (nintedanib or placebo) and received $\geq$1 dose of study medication. The annual rate of decline in FVC will be primarily analysed using a random coefficient regression (random slopes and intercepts) model including gender, age and height as covariates. All available FVC values from baseline to week 52 will be used in the primary model, including FVC measurements from the follow-up visit for patients who prematurely discontinued trial medication and did not complete study visits until week 52. A linear model was chosen as in this patient population, FVC is expected to decline linearly over time. However, a number of alternative and sensitivity analyses have been pre-specified in the statistical analysis plan, such as change from baseline to week 52 in FVC and other functional forms for the rate of decline (quadratic and exponential) to assess the robustness of the linear model. Model assumptions also include a normal distribution for the intercepts and slopes with an arbitrary covariance matrix. An unstructured variance-covariance structure will be used to model within-patient measurements. The variance-covariance matrix, modeled to estimate the inter-individual variability, will be considered to have a Variance-Components structure. The Roger-Kenward approximation will be used to estimate denominators degrees of freedom.

Change from baseline in SGRQ total score over 52 weeks will be primarily analysed using mixed model repeated measures (MMRM) with treatment and visit as fixed effects, baseline SGRQ total score as a covariate, and treatment-by-visit and baseline-by-visit as interaction terms. The patient effect will be assumed to be random and compound symmetry covariance structure will be assumed for within-patient variation.

Kaplan–Meier estimates will be derived for the probability of a first acute exacerbation over time, and time to first acute exacerbation will be primarily analysed using the log rank test. The hazard ratios and their confidence intervals will be computed using a Cox proportional hazards model adjusted for gender, age and height. These covariates were chosen in order to be consistent with the analyses performed in the Phase II TOMORROW trial [14] and are the same covariates as included in the primary endpoint model. The key secondary endpoint uses data on acute exacerbations as reported by the site investigators, in keeping with the Phase II methodology. Events adjudicated as confirmed or suspected acute exacerbations by the Adjudication Committee will be assessed in a sensitivity analysis of data pooled from both INPULSIS™ trials.

A hierarchical procedure will be used to demonstrate the superiority of nintedanib over placebo for the primary and key secondary endpoints. The consecutive steps of the hierarchy will only be considered if the previous step is significant at the 1-sided 2.5% level and the results are in favour of nintedanib. Two hierarchies of endpoints, with a different order of the key secondary endpoints for submissions to US and EU/rest of world regulatory authorities, will be tested. For the US submission, time to first acute exacerbation is the first key secondary endpoint; for the EU/rest of world submissions, change from baseline in SGRQ total score over 52 weeks is the first key secondary endpoint. No hierarchy will be used for the other secondary endpoints.

Absolute and relative changes from baseline in FVC over 52 weeks will be analysed using MMRM, with treatment and visit as fixed effects and baseline value, gender, age and height as covariates, and treatment-by-visit and baseline-by-visit as interaction terms. Changes in other respiratory parameters will be analysed in the same way as change in FVC. Changes in other PROs will be analysed in the same way as change in SGRQ total score.

For the survival analyses, a log rank test will be used to compare treatment groups and a Cox model adjusted for gender, age and height will be used to determine hazard ratios. These covariates were chosen in order to be consistent with the analyses performed in the Phase II TOMORROW trial [14] and are the same covariates as included in the primary endpoint model. Since the number of deaths is expected to be low, the protocol specified that survival analyses will additionally be performed on the pooled data from both INPULSIS™ trials. Safety analyses will be descriptive.

Sensitivity analyses will be performed to assess the robustness of the results of the primary and key secondary endpoints. Model assumptions will be checked and sensitivity to data handling, including missing data handling, will be assessed. In order to improve the precision of the treatment effect estimates for the efficacy endpoints and to increase the size of the safety database, a pooled analysis of the two trials was pre-specified as an additional supportive analysis.

## Discussion

### Rationale for dose selection

The dose of nintedanib used in the INPULSIS™ trials was selected based on findings from the 12-month Phase II TOMORROW study [14]. In the TOMORROW trial, the annual rate of decline in FVC in the nintedanib 150 mg twice daily group was −0.06 L (95% CI, −0.14 to 0.02) compared with −0.19 L (95% CI, −0.26 to −0.12) in the placebo group: a difference of 0.13 L (95% CI, 0.03–0.24). In addition, treatment with nintedanib 150 mg twice daily was associated with preservation of HRQoL (mean change in SGRQ total score of −0.66 [95% CI, −4.02 to 2.71] versus 5.46 [95% CI, 2.06, 8.86] with placebo: a difference of −6.12 [95% CI, −10.57 to −1.67]) and a reduction in the risk of acute exacerbations (risk ratio compared with placebo: 0.16 [95% CI, 0.03 to 0.70]).

### Rationale for endpoints

The most robust primary endpoint for Phase III clinical trials in IPF is all-cause mortality [25]. However, the mortality rate of patients enrolled in the TOMORROW trial was low,