Visual analogue scales for interstitial lung disease: a prospective validation study

Helen Yates, Huzaifa I Adamali, Nick Maskell, Shaney L Barratt, Charles Sharp

Miss Helen Yates, MSc – Respiratory Physiology, North Bristol NHS Trust, Bristol BS10 5NB, UK

Dr Huzaifa I Adamali, MD – Bristol ILD Service, North Bristol NHS Trust, Bristol BS10 5NB, UK

Prof Nick Maskell, MD – Academic Respiratory Unit, University of Bristol, Southmead Hospital, Bristol BS10 5NB, UK

Dr Shaney L Barratt, PhD – Bristol ILD Service, North Bristol NHS Trust, Bristol BS10 5NB, UK

Dr Charles Sharp, MD – Respiratory Department, Gloucestershire Hospitals NHS Foundation Trust, Gloucester Royal Hospital, Great Western Road, Gloucester GL1 3NN, UK

Corresponding Author – Charles Sharp – Charles.sharp@doctors.org.uk

Word count

All authors declare no conflicts of interest.

Abbreviations

6MWT – 6-minute walk test
ANOVA – Analysis of variance
ASS – Anti-synthetase syndrome
CI – Confidence interval
CPFE – Combined pulmonary fibrosis and emphysema
DLCO – Diffusing capacity for Carbon Monoxide
ERES – Empirical rule effect size
ES – Effect size
FVC – Forced vital capacity
HP – Hypersensitivity pneumonitis
HRQL – Health-related quality of life
ILD – Interstitial lung disease
IPF – Idiopathic pulmonary fibrosis
K-BILD – King's Brief ILD questionnaire
MCID – Minimal clinically important difference
NHS – National Health Service
NSIP – Non-specific interstitial pneumonia
PLCH – Pulmonary Langerhans cell Histiocytosis
RA-UIP – Rheumatoid Arthritis-associated usual interstitial pneumonia
VAS – Visual analogue scale
VASC – VAS for cough
VASD – VAS for dyspnoea
VASF – VAS for fatigue

ABSTRACT

Background Visual analogue scales (VAS) are simple symptom assessment tools which have not been validated in interstitial lung disease (ILD). Simple measures of ILD disease burden would be valuable for non-specialist clinicians monitoring disease away from ILD specialist centres. This study aimed to validate VAS to assess change in dyspnoea, cough and fatigue in ILD, and to define the minimal clinically important difference (MCID) for change in these.

Methods 64 patients with ILD completed VAS for dyspnoea, cough and fatigue. Baseline King's Brief ILD questionnaire (K-BILD) scores, lung function and 6-minute walk test results were collected. Tests were repeated 3-6 months later, in addition to a 7-point Likert scale. The MCID was estimated using median change in VAS in patients who reported "small but just worthwhile change" in symptoms at follow-up. Methods were repeated in a validation cohort of 31 ILD patients to confirm findings.

Results VAS scores were significantly higher for patients who reported a "small but just worthwhile change" in symptoms versus "no change" or "not worthwhile change" ($p<0.01$). The MCID for VAS Dyspnoea was estimated as 22.0mm and 14.5mm for VAS Fatigue. These results were reproducible in the validation cohort. Results were not significant for VAS Cough. Change in VAS Dyspnoea correlated with change in K-BILD ($r=-0.51$, $p<0.01$), forced vital capacity ($r=-0.32$, $p=0.01$) and 6-minute walking distance ($r=-0.37$, $p=0.01$).

Conclusion: The VAS is valid for assessing change in dyspnoea and fatigue in ILD. The MCID is estimated as 22.0mm for dyspnoea and 14.5mm for fatigue. This could be used to monitor disease in settings away from ILD specialist review.

INTRODUCTION

Interstitial lung diseases (ILDs) are chronic, progressive disorders of the lung parenchyma associated with significant morbidity and mortality[1]. Despite recent advances in ILD therapies[3], the care of a large proportion of ILD patients focuses on the management of dyspnoea, cough and fatigue[2,4,5]. Objective measures of lung function and exercise testing may not reflect patient experience of disease, and HRQL tools can be time consuming to complete and interpret, limiting their use.

Increasingly, the model of care for patients with ILD is of shared care between a local, referring centre and a specialist ILD centre. This care may incorporate community nurses, non-specialist physicians and other allied health professionals. Simple, quickly completed tools to assess patients' symptoms in ILD would be valuable additions to clinical assessment outside well-resourced specialist centres[2].

The visual analogue scale (VAS) is a simple tool for assessing patients' symptoms, which has been validated in asthma[6], chronic obstructive pulmonary disease[7] and pleural disease[8]. Only one study has assessed VAS in ILD, however the sample size was small (n=27) and this did not assess change over time[9].

The aim of this study was to investigate whether VAS are valid tools for assessing change in dyspnoea, cough and fatigue in a diverse range of ILD, as encountered in a heterogeneous ILD clinic, and to identify the minimal change considered worthwhile to the patient[10], also known as the minimum clinically important difference (MCID)[11].

METHODS

Assessing validity of VAS and determining the MCID

This was a prospective observational study. Ethical approval was granted by East of Scotland Research Ethics Service and all participants provided written consent. Patients with a range of ILDs were recruited consecutively from the Bristol ILD service at North Bristol NHS Trust from January-November 2016. Inclusion criteria were: a diagnosis of ILD, English speaking, and a minimum of primary school education.

Patients completed written VAS for dyspnoea, cough and fatigue, lung function tests, a 6-minute walk test (6MWT) and the Kings Brief ILD (K-BILD) HRQL tool. VAS design was a 100mm

continuous, horizontal line labelled to express symptom extremes, based on existing VAS for other respiratory disease (Figure 1)[12]. These assessments were completed with the assistance of respiratory physiologists supervising their lung function testing. At follow-up after 3-6 months, patients completed a second VAS, a 7-point Likert scale, lung function tests, 6MWT and repeated the K-BILD tool (Figure 1). Patients were shown their previous VAS scores as this has been demonstrated to increase the validity of patient reported outcomes[13].

VAS scores were calculated by measuring the distance in millimetres from the start of the line to the centre of the point recorded by the patient[12]. Change in VAS was calculated from the difference between VAS score at initial and follow-up visits. In analysing results, the relative change in symptoms was explored by combining improvement and deterioration together. This resulted in four categories for analysis. Data were categorised according to Likert scale response; "no change", "slight change but not worthwhile", "small but just worthwhile change", and "large or moderate change" (Figure 1). The MCID was estimated using the median change in VAS in the "small but just worthwhile change" category.

Lung function and 6MWT results were selected as clinical anchors to determine the MCID[14]. These are validated measures with described MCID, used in prognostication and monitoring for ILD[15,18,19]. Patient-based anchors were the 7-point Likert scale[14,20] and the K-BILD tool[17].

Estimates of MCID were compared to distributional methods including effect size (ES) and empirical rule effect size (ERES)[21]. Distributional methods provide statistical estimates of MCID from underlying variation within the sample[21]. An ES of 0.33 has been suggested to equate to the MCID[22], therefore ES MCID was estimated by multiplying the standard deviation of baseline VAS score by 0.33. The ERES is based on the assumption that the mean score is half of the maximum score (in this case 100mm), and the range of the outcome measure is 6 standard deviations[21]. Therefore the ERES MCID was calculated by dividing 100 by 6. Additionally, patients estimated the change in VAS which they considered "meaningful" to provide patient-opinion estimated MCID. Results were compared with anchor and distribution-based estimates.

Statistical Analysis

Statistical analyses were performed using MiniTab17 Statistical software[23]. All patients who completed data collection at follow-up were included in the analyses. Data were assessed for normality and one way analysis of variance (ANOVA) was used to assess for differences between

groups categorised by Likert score. Moods Median test was used for non-parametric data. Median change in VAS for patients reporting a "small but just worthwhile change" was used to calculate MCID.

Correlations between change in VAS and change in forced vital capacity (FVC), diffusing capacity for carbon monoxide (DL_{CO}), 6MWT and K-BILD scores were assessed using Spearman's Rank correlation coefficient. The strength of correlations were determined according to absolute values of the coefficient; large (r>0.5), moderate (r=0.5-0.3) and small (r=0.1-0.3)[24]. Any influence of symptom severity on patients' perceived change in symptoms was assessed by examining correlation between baseline VAS and subsequent change in VAS.

A sample size of 8 patients in the Likert category "small but just worthwhile change" was required to give a power of 90% to detect a change in VAS scores at a p-value of 0.05, calculated from the MCID and standard deviation reported in a study of dyspnoea related to pleural disease[8,25]. The validity of the MCID was assessed by repeating these methods in a second cohort. The estimated MCID was assessed for similarity with the initial cohort using the Mann-Whitney U test. The recruitment cohort sizes were pragmatically selected to maximise the chances of achieving this in the absence of published data in this area for ILD.

Outcomes

The primary outcome of the study was the MCID estimated from the median change in VAS for patients reporting a "small but just worthwhile change" on the 7-point Likert scale at follow-up. Other pre-specified secondary outcomes included correlation of change in VAS with change in FVC, DL_{CO}, 6MWT and K-BILD, and comparison of the estimated MCID with patient-opinion and distributional methods**.**

RESULTS

Validating VAS and determining the MCID

Of 131 patients recruited for this study, one was excluded due to visual impairment. Completed data were available for 95/130 (73%) of patients enrolled (2 died, 33 did not attend for follow-up during the study period). The first 64 completed data sets were assigned to the initial cohort, and subsequent 31 to the validation cohort (Figure 2).

The baseline characteristics of patients in both cohorts were comparable and are shown in Table 1. Mean age in the initial cohort was 66 years, 41% were female, 97% were Caucasian, and 31% had Idiopathic Pulmonary Fibrosis (IPF). Mean FVC was 83% predicted and mean DL_{CO} was 55.7% predicted.

Overall, there were 30 patients with IPF and 67 with non-IPF fibrotic lung disease included in the analyses. There was no difference between these groups for baseline or interval change in VASD, VASC or VASF (Supplementary Table 1). There was no difference in the proportion of respondents reporting "small but just worthwhile change" in symptoms who had IPF when compared to the proportion of respondents with IPF in the overall cohort (38.6% vs 30.9%, p=0.304).

Anchor-based MCID

Likert scale selections were similarly distributed for dyspnoea, fatigue and cough (Table 2). Most patients reported "no change" in symptoms whilst the lowest numbers reported "moderate or large change". Similar numbers reported a "small but just worthwhile change" for all symptoms.

Median changes in VAS categorised by Likert scale response are shown in Table 2. Median change in VAS increased as Likert response increased for all symptoms. The MCID for change in VAS Dyspnoea (VASD) was 22.0mm, equating to a "small but just worthwhile change" on the Likert scale. Moods median showed statistically significant differences in change in VASD between all Likert groups ($p<0.001$, 95% CI, 12-35mm) (Table 2, Figure 3A).

The MCID for change in VAS Fatigue (VASF) was 14.5mm. There were statistically significant differences for change in VASF between all Likert groups (p=0.006, 95% CI, 8-20mm) (Table 2, Figure 3B). There were no significant between-group differences in change in VAS Cough (VASC) (p=0.061) (Figure 3C), therefore the MCID could not be determined.

Distribution-based estimates of MCID were lower than anchor-based estimates for VASD (7.7mm-8.4mm and 22.0mm respectively), but values were similar for VASF (8.4mm- 9.0mm and 14.5mm respectively) (Table 3). Patient-opinion and anchor-based estimates of VASD MCID were similar (20.5mm and 22.0mm respectively), but patient-opinion was higher for VASF (28.0mm and 14.5mm) (Table 3).

Correlations

Change in VASD correlated moderately with change in FVC% (r=-0.319) and 6MWD (r=-0.365), but there was no correlation with change in DL_{CO}% (Table 4). The correlations between change in VASF and 6MWD were moderate (r=-0.349), but insignificant for FVC% and DL_{CO}% (Table 4). Change in VASD, VASF and VASC all moderately correlated with K-BILD scores (r=-0.363—0.506). The strongest correlation was between VASD and dyspnoea-specific K-BILD domains (r=-0.557) (Figure 4). There was no correlation between VASC and K-BILD cough domains, lung function or 6MWD (Table 4). There was no correlation between any VAS and age.

There was no correlation between initial VASD and VASF and subsequent change in VAS (r=-0.075 and r=-0.184 respectively). There was moderate correlation between VASC and subsequent change in VAS (r=-0.364, p<0.01).

Validation of the MCID

Baseline characteristics for the validation cohort of 31 patients are shown in Table 1. Anchor-based estimates of the MCID were 26.5mm for change in VASD and 11.0mm for change in VASF, for the validation cohort (Table 3). These were not significantly different for both VASD (p=0.66, 95% CI, -19-13mm) and VASF (p=0.77, 95% CI, -7-16mm).

DISCUSSION

This is the first study to demonstrate that VAS are valid tools for assessing change in dyspnoea and fatigue in ILD. The MCID for change in VASD is 22.0mm and VASF is 14.5mm. VAS are simple, quick and easy to use patient-reported measures, which could have value in clinical assessment outside specialist ILD centres. The validity of VAS and their responsiveness to change were shown by correlation with validated measures of disease status.

The authors recognise that there are some limitations, including those inherent to a single centre study, therefore caution should be taken before applying these findings to other populations. 33 patients did not return for follow-up during the study period due to the regional status of the ILD service patient. Additionally, this work was conducted in a heterogeneous cohort, including some patients with sarcoidosis, which may have influenced the findings of the VASF. These patients all had fibrotic parenchymal disease, however and the cohort was selected to represent a pragmatic, real-world clinical spectrum reviewed in ILD centres. The grouping together of IPF, a progressive disease, with non-IPF fibrotic lung diseases has the potential to influence these results, however no statistically significant differences were seen between the VAS for these groups at baseline or

on follow-up. Likewise, IPF was not over-represented in the group of respondents reporting "small but just worthwhile change" in symptoms at follow-up.

The selection of clinical parameters to which to compare the VAS was based on those used routinely in this ILD centre. As such, we did not compare VASC and VASF to specific cough and fatigue questionnaires. This may limit the interpretation of our results, however the VASC did not reveal significant changes over time. The K-BILD QoL tool, while not specifically designed to assess fatigue, does give a holistic assessment of patient symptoms and as such is an appropriate comparator to the VASF and VASD. There are also limitations inherent to the use of VAS. These measures are subject to "end of scale" bias, wherein respondents are less likely to use the extreme ends of the scale to assess their health status. Likewise, it is possible that while VAS are measured to the nearest millimetre, respondents are unable to make such fine distinctions in position along a line. This should not prevent their application as a simple tool, easily interpreted by clinicians without experience or expertise in specialist QoL tools.

Dyspnoea has been reported as the principal and most debilitating symptom in ILD[4], therefore its assessment is essential in disease management. The estimated MCID of 22.0mm, confirmed in a second cohort of 31 patients, is similar to reports of dyspnoea in pleural disease[8], asthma[6] and COPD[7]. VASD demonstrated responsiveness to change in ILD status by correlation with changes in validated measures, including FVC, 6MWD and K-BILD[15,27]. In contrast, change in VASD did not correlate with change in DL_{CO}%. Although DL_{CO}% is used to describe IPF severity and predict disease progression[18], its variability compared with FVC[28] limits its use as a primary outcome in IPF clinical trials[29,30].

Dyspnoea is influenced by several factors in addition to the pulmonary restriction and reduced gas exchange observed in ILD, and the relationships between dyspnoea and FVC, DLCO and 6MWD are complex[2,27]. Previous studies have found symptom domains of quality of life assessments did not correlate with FVC% and DL_{CO}%[31,32], whereas other studies have found weak but statistically significant associations[1,19,33]. A clinical trial in patients with IPF found significant reductions in FVC were not associated with reduced dyspnoea scores[29]. A possible explanation could be that patients with more severe disease status limit their activity levels and therefore under-report symptom deterioration[31]. This is reflected in the low mean values for VASD observed in this study (34mm).

Our findings show VASD is also associated with validated patient-reported measures. Change in VASD score was associated with change in Likert score, demonstrating its ability to reflect patient experience[14]. Furthermore change in VASD correlated with change in K-BILD scores, consistent with previous studies which found correlations between VASD and other quality of life assessments (r=-0.61)[33]. Finally, patient-estimated MCID was similar to anchor-based methods, providing further evidence that VASD accurately reflects patient opinion.

Distributional methods underestimated the MCID for VASD, consistent with a study which also compared distributional and anchor-based methods[8]. Distributional approaches have been criticised as they do not use clinical anchors and provide a purely statistical estimation of MCID based on the underlying variation within the sample[3,4,8,14]. In this study we used the optimal approach recommended for determining the MCID, with increased emphasis on anchor methods, and support provided by distributional methods and patient opinion[14,17].

The results of our study suggest that VAS is valid for assessing change in fatigue in patients with ILD, and the estimated MCID is 14.5mm. Median VAS fatigue (VASF) score was higher for patients reporting a “small but just worthwhile change” compared to “no change”, and median change in VASF increased in association with increased Likert scale response. Additionally, the MCID was confirmed in the validation cohort. However, there were wide and overlapping 95% CI for all groups, particularly those reporting changes, and a lack of correlation of VASF with other measures. These results are consistent with a study which found VASF was less sensitive than other scales for assessing fatigue, whereas VASD was superior[35].

There is a paucity of studies which have looked specifically at the prevalence of fatigue in ILD, although it is widely understood to be common based on quality of life studies[5]. Predictors of fatigue have been modelled for IPF but could not be identified for sarcoidosis, indicating the diverse manifestation of the symptom amongst different ILDs[36]. Although VASF has not been investigated in ILD, it has been assessed in a range of diseases including rheumatoid arthiritis[37] and cancer[38] and was validated for patients with sleep disorders[39]. VASF has also been shown to correlate with objective measures including exercise testing. However, the study populations used were small and included male athletes[35] and Chinese female students[40], limiting their applicability to our study population.

The correlation observed between VASF and 6MWD could be due to the inclusion of physical muscle weakness secondary to de-conditioning within the subjective perception of fatigue[5]. Therefore although VASF was able to detect change in fatigue, given the complexity of fatigue symptoms and the mixed results observed in this study, further research is required before VASF can be recommended for clinical use in patients with ILD.

In this study VASC was unable to detect change in cough symptoms. Previous studies have demonstrated a lack of correlation between cough rates and lung function tests[41] and between subjective and objective measures of cough[42,43]. Cough is a complex symptom of ILD, and there are likely to be alternative causes such as rhinitis and gastro-oesophageal reflux in the patients studied[44]. Patients have identified cough as a fundamental symptom of ILD[45], therefore it is important to establish valid methods for its assessment.

This is the first study to assess change in VASD, VASF and VASC in ILD. The sample size was larger than previous studies, and longitudinal study design allowed for responsiveness of VAS to be determined, and calculated MCIDs were validated in a separate cohort. This study was performed in a busy out-patient service and demonstrates the practicality of the VAS for routine use in ILD clinics.

Conclusion

This study has shown the VAS is a valid and clinically relevant tool for assessing change in dyspnoea and fatigue in patients with ILD. These scales correlate with recognised markers of disease, including K-BILD and pulmonary physiology tests. The MCID for change in VAS is 22.0mm for dyspnoea and 14.5mm for fatigue. VAS is a quick and simple tool which could be used alongside lung function tests and quality of life assessments to establish disease progression and could represent a useful adjunct to clinical assessment in non-specialist settings where more time-consuming measures are not practical.

Acknowledgements

This work was conceived by CS, NM and HY. Patients were recruited by HY, SB and HA. Data analysis was conducted by HY and CS. The manuscript was composed by HY and CS and all authors approved the final text. CS is guarantor for this work.

We would like to thank patients who participated in the study, Jason Viner and the Respiratory Physiology department at North Bristol NHS Trust for their help in identifying patients and the Bristol ILD Service team for supporting this work.

REFERENCES

1. Patel AS, Siegert RJ, Brignall K, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. *Thorax* 2012;67(9):804–810.
2. Bajwah S, Ross JR, Peacock JL, et al. Interventions to improve symptoms and quality of life of patients with fibrotic interstitial lung disease: a systematic review of the literature. *Thorax* 2013;68(9):867–879.
3. Belloli EA, Martinez FJ, Flaherty KR. Update in Interstitial Lung Disease 2014. *Am J Respir Crit Care Med* 2015;192(5):538–543.
4. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *Am J Respir Crit Care Med* 2000;161(2 Pt 1):646–664.
5. Garibaldi BT, Danoff SK. Symptom-based management of the idiopathic interstitial pneumonia. *Respirol Carlton Vic* 2016;21(8):1357–1365.
6. Karras DJ, Sammon ME, Terregino CA, Lopez BL, Griswold SK, Arnold GK. Clinically meaningful changes in quantitative measures of asthma severity. *Acad Emerg Med Off J Soc Acad Emerg Med* 2000;7(4):327–334.
7. Torres JP de, Pinto-Plata V, Ingenito E, et al. Power of outcome measurements to detect clinically significant changes in pulmonary rehabilitation of patients with COPD. *Chest* 2002;121(4):1092–1098.
8. Mishra EK, Corcoran JP, Hallifax RJ, Stradling J, Maskell NA, Rahman NM. Defining the minimal important difference for the visual analogue scale assessing dyspnea in patients with malignant pleural effusions. *PloS One* 2015;10(4):e0123798.
9. Ramírez-Venegas A, Sansores Martínez RH, Carrillo Rodríguez G, Salas Hernández J, Chapela Mendoza R, Selman Lama M. [Validation of a visual analog scale to measure dyspnea in patients with diffuse interstitial lung disease]. *Rev Investig Clínica Organo Hosp Enfermedades Nutr* 1994;46(6):479–486.
10. Redelmeier DA, Lorig K. Assessing the clinical importance of symptomatic improvements. An illustration in rheumatology. *Arch Intern Med* 1993;153(11):1337–1342.
11. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10(4):407–415.
12. Ries AL. Minimally Clinically Important Difference for the UCSD Shortness of Breath Questionnaire, Borg Scale, and Visual Analog Scale. *COPD J Chronic Obstr Pulm Dis* 2005;2(1):105–110.
13. Guyatt GH, Townsend M, Keller JL, Singer J. Should study subjects see their previous responses: data from a randomized control trial. *J Clin Epidemiol* 1989;42(9):913–920.
14. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61(2):102–109.
15. Bois RM du, Weycker D, Albera C, et al. Forced Vital Capacity in Patients with Idiopathic Pulmonary Fibrosis: Test Properties and Minimal Clinically Important Difference. *Am J Respir Crit Care Med* 2011;184(12):1382–1389.
16. Bois RM du, Weycker D, Albera C, et al. Six-Minute-Walk Test in Idiopathic Pulmonary Fibrosis: Test Validation and Minimal Clinically Important Difference. *Am J Respir Crit Care Med* 2011;183(9):1231–1237.
17. Patel AS, Siegert RJ, Keir GJ, et al. The minimal important difference of the King's Brief Interstitial Lung Disease Questionnaire (K-BILD) and forced vital capacity in interstitial lung disease. *Respir Med* 2013;107(9):1438–1443.
18. Meyer KC. Diagnosis and management of interstitial lung disease. *Transl Respir Med* 2014;2:4.
19. Albera C. Challenges in idiopathic pulmonary fibrosis trials: the point on end-points. *Eur Respir Rev* 2011;20(121):195–200.
20. Ander DS, Aisiku IP, Ratcliff JJ, Todd KH, Gotsch K. Measuring the dyspnea of decompensated heart failure with a visual analog scale: how much improvement is meaningful? *Congest Heart Fail Greenwich Conn* 2004;10(4):188–191.

321. Sloan JA, Dueck A. Issues for Statisticians in Conducting Analyses and Translating Results for Quality of Life
324 End Points in Clinical Trials. *J Biopharm Stat* 2004;14(1):73–96.
325 2. Yost KJ, Cella D, Chawla A, et al. Minimally important differences were estimated for the Functional
326 Assessment of Cancer Therapy–Colorectal (FACT-C) instrument using a combination of distribution- and
327 anchor-based approaches. *J Clin Epidemiol* 2005;58(12):1241–1251.
328 3. Minitab 17 Statistical Software. [Computer software]. State College, PA: Minitab, Inc; 2010.
329 4. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2 edition. Hillsdale, N.J: Routledge; 1988.
330 5. Lehr R. Sixteen S-squared over D-squared: a relation for crude sample size estimates. *Stat Med*
331 1992;11(8):1099–1102.
332 6. Anthoine E, Moret L, Regnault A, Sébille V, Hardouin J-B. Sample size used to validate a scale: a review of
333 publications on newly-developed patient reported outcomes measures. *Health Qual Life Outcomes*
334 [Internet] 2014 [cited 2016 Jun 16];12(1). Available from:
335 http://hqlo.biomedcentral.com/articles/10.1186/s12955-014-0176-2
336 7. Swigris JJ, Han M, Vij R, et al. The UCSD shortness of breath questionnaire has longitudinal construct validity
337 in idiopathic pulmonary fibrosis. *Respir Med* 2012;106(10):1447–1455.
338 8. Berend N. Respiratory disease and respiratory physiology: putting lung function into perspective interstitial
339 lung disease. *Respirol Carlton Vic* 2014;19(7):952–959.
340 9. King TE, Bradford WZ, Castro-Bernardini S, et al. A Phase 3 Trial of Pirfenidone in Patients with Idiopathic
341 Pulmonary Fibrosis. *N Engl J Med* 2014;370(22):2083–2092.
342 10. Richeldi L, Bois RM du, Raghu G, et al. Efficacy and Safety of Nintedanib in Idiopathic Pulmonary Fibrosis. *N*
343 *Engl J Med* 2014;370(22):2071–2082.
344 11. Chang JA, Curtis JR, Patrick DL, Raghu G. Assessment of health-related quality of life in patients with
345 interstitial lung disease. *Chest* 1999;116(5):1175–1182.
346 12. Swigris JJ, Yorke J, Sprunger DB, et al. Assessing dyspnea and its impact on patients with connective tissue
347 disease-related interstitial lung disease. *Respir Med* 2010;104(9):1350–1355.
348 13. Khanna D, Clements PJ, Furst DE, et al. Correlation of the degree of dyspnea with health-related quality of
349 life, functional abilities, and diffusing capacity for carbon monoxide in patients with systemic sclerosis and
350 active alveolitis: Results from the scleroderma lung study. *Arthritis Rheum* 2005;52(2):592–600.
351 14. Sloan JA, Frost MH, Berzon R, et al. The clinical significance of quality of life assessments in oncology: a
352 summary for clinicians. *Support Care Cancer* 2006;14(10):988–998.
353 15. Grant S, Aitchison T, Henderson E, et al. A comparison of the reproducibility and the sensitivity to change of
354 visual analogue scales, Borg scales, and Likert scales in normal subjects during submaximal exercise. *Chest*
355 1999;116(5):1208–1217.
356 16. Atkins CP, Gilbert D, Brockwell C, Robinson S, Wilson AM. Fatigue in sarcoidosis and idiopathic pulmonary
357 fibrosis: differences in character and severity between diseases. *Sarcoidosis Vasc Diffuse Lung Dis Off J*
358 *WASOG* 2016;33(2):130–138.
359 17. Bala S-V, Forslind K, Månsson ME. Self-reported outcomes during treatment with tumour necrosis factor
360 alpha inhibitors in patients with rheumatoid arthritis. *Musculoskeletal Care* 2010;8(1):27–35.
361 18. Baró E, Carulla J, Cassinello J, et al. Development of a new questionnaire to assess patient perceptions of
362 cancer-related fatigue: item generation and item reduction. *Value Health J Int Soc Pharmacoeconomics*
363 *Outcomes Res* 2009;12(1):130–138.
364 19. Lee KA, Hicks G, Nino-Murcia G. Validity and reliability of a scale to assess fatigue. *Psychiatry Res*
365 1991;36(3):291–298.
366 20. Leung AWS, Chan CCH, Lee AHS, Lam KWH. Visual analogue scale correlates of musculoskeletal fatigue.
367 *Percept Mot Skills* 2004;99(1):235–246.
368 21. Key AL, Holt K, Hamilton A, Smith JA, Earis JE. Objective cough frequency in Idiopathic Pulmonary Fibrosis.
369 *Cough Lond Engl* 2010;6:4.
370 22. Hamutcu R, Francis J, Karakoc F, Bush A. Objective monitoring of cough in children with cystic fibrosis.
371 *Pediatr Pulmonol* 2002;34(5):331–335.
372 23. Leconte S, Ferrant D, Dory V, Degryse J. Validated methods of cough assessment: a systematic review of the
373 literature. *Respir Int Rev Thorac Dis* 2011;81(2):161–174.
374 24. Lavorini F, Spina D, Walker MJ, Franciosi L, Page CP, Fontana GA. Antitussive effect of carcainium chloride in
375 patients with chronic cough and idiopathic interstitial pneumonias: A pilot study. *Pulm Pharmacol Ther*
376 2016;40:91–94.

375. Saketkoo LA, Mittoo S, Huscher D, et al. Connective tissue disease related interstitial lung diseases and idiopathic pulmonary fibrosis: provisional core sets of domains and instruments for use in clinical trials. *Thorax* 2014;69(5):428–436.

Tables

Table 1 Baseline characteristics by cohort

	Initial cohort (N=64)	Validation Cohort (N=31)
Mean age, years (SD)	66 (14)	68 (12)
Mean follow-up, days (SD)	127 (58)	116 (37)
Women, n (%)	26 (41)	13 (42)
Race/ethnicity, n (%)		
Caucasian	62 (97)	31 (100)
Other*	2 (3)	0 (0)
Smoking status, n (%)		
Never smoked	31 (48)	16 (52)
Ex-smoker	31 (48)	15 (48)
Current smoker	2 (3)	0 (0)
Mean baseline VAS score, mm (SD)		
VAS Dyspnoea	34 (23)	35 (24)
VAS Cough	43 (26)	41 (30)
VAS Fatigue	43 (27)	40 (25)
Mean baseline lung function tests		
FVC, L (SD)	2.8 (1.0)	2.9 (1.2)
FVC, % predicted (SD)	82.5 (18.8)	88.9 (20.1)
DLCO, % predicted (SD)	55.7 (19.6)	56.6 (22.4)
6MWD, m (SD)	340 (111.6)	320.8 (96.0)
Nadir oxygen saturations at 6MWT, % (SD)	87.6 (5.8)	88 (5.3)
Mean baseline K-BILD score (SD)	62.6 (21.4)	62.5 (22.7)
ILD diagnosis, n (%)		
IPF	20 (31)	10 (32)
HP	9 (14)	7 (23)
Sarcoidosis (fibrotic parenchymal disease)	8 (13)	4 (13)
CT-ILD	7 (11)	3 (10)
NSIP	5 (8)	3 (10)
Other	10 (16)	4 (13)
Unclassifiable	5 (8)	0 (0)

Definition of abbreviations: FVC, forced vital capacity; DLCO, diffusing capacity for carbon monoxide; 6MWD, 6-minute walk distance; K-BILD, King's Brief Interstitial Lung Disease questionnaire; CT-ILD, Connective Tissue Disease-associated ILD; HP, Hypersensitivity Pneumonitis; IPF, Idiopathic pulmonary fibrosis; NSIP, non-specific interstitial pneumonia.

*South Asian patients originating from India, Pakistan or Bangladesh

391

392

393

394

395

Table 2 Median change in VAS categorised by Likert scale response

	Likert Scale Response	Number of patients (%)	Median change VAS (mm)	95% CI (mm)
VAS Dyspnoea	No change	25 (39)	4.0	2-6
	Slight change but not worthwhile	16 (25)	9.0	6-10
	Small but just worthwhile change*	17 (27)	22.0	12-35
	Large or moderate change	6 (9)	30.0	12-64
VAS Fatigue	No change	17 (27)	4.0	2-9
	Slight change but not worthwhile	19 (30)	11.0	5-17
	Small but just worthwhile change*	18 (28)	14.5	8-20
	Large or moderate change	10 (16)	20.5	6-49
VAS Cough	No change	18 (28)	7.0	3-18
	Slight change but not worthwhile	19 (30)	10.0	7-17
	Small but just worthwhile change*	18 (28)	18.0	15-32
	Large or moderate change	9 (14)	23.0	11-77

Definition of abbreviations: VAS, visual analogue scale; CI, confidence interval.
*Median change in VAS score for patients who reported a "small but just worthwhile change" in symptoms on the Likert scale was used to estimate the MCID

396

397

Table 3 Anchor, Distributional and Patient-opinion estimates of the minimal clinically important difference (MCID)*

	ES	ERES	Anchor-based	Patient- opinion
VAS Dyspnoea	7.7	8.4	22.0	20.5
VAS Fatigue	9.0	8.4	14.5	28.0

N=64
Definition of abbreviations: MCID, minimal clinically important difference; ES, effect size; ERES, empirical rule effect size
*MCID in mm

398

399

Table 4 Spearman's correlation coefficients between change in VAS and change in other measures

	VAS Dyspnoea	VAS Fatigue	VAS Cough
FVC%	-0.319 (p=0.010)	-0.275 (p=0.028)	0.06 (p=0.635)
DLCO%	-0.201 (p=0.124)	-0.186 (p=0.156)	-0.012 (p=0.928)
6MWD	-0.365 (p=0.007)	-0.349 (p=0.010)	0.045 (p=0.751)
KBILD Overall	-0.506 (p=0.000)	-0.500 (p=0.000)	-0.363 (p=0.003)
KBILD Specific	-0.557 (p=0.000)	-0.423 (p=0.000)	-0.217 (p=0.085)

N=64
Definition of abbreviations: VAS, visual analogue scale; FVC, forced vital capacity; DLCO, carbon monoxide diffusing capacity; 6MWD, 6-minute walk distance; K-BILD, King's Brief Interstitial Lung Disease Questionnaire.

400

Supplementary Table 1
Comparison of IPF and non-IPF fibrotic lung disease VAS at baseline and interval change

Visual Analogue Scales in ILD

	IPF	Non-IPF	p-value
Mean baseline VAS Dyspnoea, mm (SD)	36 (19)	34 (25)	0.459
Mean baseline VAS Cough, mm (SD)	38 (20)	44 (29)	0.490
Mean baseline VAS Fatigue, mm (SD)	43 (21)	42 (29)	0.674
Change in VAS Dyspnoea, mm (SD)	3 (17)	2 (21)	0.603
Change in VAS Cough, mm (SD)	2 (20)	-5 (29)	0.298
Change in VAS Fatigue, mm (SD)	9 (19)	2 (21)	0.124

Definition of abbreviations: VAS, visual analogue scale; IPF, Idiopathic Pulmonary Fibrosis.

401

Figure Legends

Figure 1 Visual Analogue Scales and 7-point Likert scale used to assess change in dyspnoea, cough and fatigue.

Figure 2 Study flowchart

Data collection was performed between January-November 2016. 33 patients who did not return for follow-up during this period were excluded from analyses

Figure 3 Comparison of change in VAS in patients categorised by Likert scale response.

The boxes show the median values and 25-75% percentiles, and lines show 0-90% percentiles. A) Change in VAS dyspnoea, significant between-group difference (p=0.000). B) Change in VAS Fatigue, significant between-group difference (p=0.006) C) Change in VAS Cough, no significant difference between groups (p=0.061). Definition of abbreviations: VAS, visual analogue scale.

Figure 4 Relationship between change in VAS Dyspnoea and dyspnoea-specific components of the K-BILD questionnaire.

Circles represent individual data points (change VASD and change in dyspnoea components of K-BILD). Solid like represents the line of best fit.

*Spearman Rank correlation coefficient

Definition of abbreviations: VAS, visual analogue scale; K-BILD, King's Brief Interstitial Lung Disease Questionnaire.