

# Responsiveness and meaningful change thresholds of the Living with Pulmonary Fibrosis (L-PF) questionnaire Dyspnoea and Cough scores in patients with progressive fibrosing interstitial lung diseases

Jeffrey J Swigris ,<sup>1</sup> Donald M Bushnell,<sup>2</sup> Klaus Rohr,<sup>3</sup> Heiko Mueller,<sup>4</sup> Michael Baldwin,<sup>3</sup> Yoshikazu Inoue <sup>5</sup>

**To cite:** Swigris JJ, Bushnell DM, Rohr K, et al. Responsiveness and meaningful change thresholds of the Living with Pulmonary Fibrosis (L-PF) questionnaire Dyspnoea and Cough scores in patients with progressive fibrosing interstitial lung diseases. *BMJ Open Res* 2022;**9**:e001167. doi:10.1136/bmjresp-2021-001167

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjresp-2021-001167>).

Received 8 December 2021  
Accepted 12 February 2022



© Author(s) (or their employer(s)) 2022. Re-use permitted under CC BY-NC. No commercial re-use. See rights and permissions. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Professor Jeffrey J Swigris;  
swigrisj@njhealth.org

## ABSTRACT

**Background** The Living with Pulmonary Fibrosis (L-PF) questionnaire assesses symptoms and quality of life in patients with fibrosing interstitial lung diseases (ILDs). Its Dyspnoea and Cough domains, whose items' responses are based on a 24-hour recall, have scores ranging from 0 to 100, with higher scores indicating greater symptom severity. We evaluated the ability of these domain scores to detect change and estimated their meaningful change thresholds in patients with progressive fibrosing ILDs.

**Methods** The INBUILD trial enrolled subjects with progressive fibrosing ILDs other than idiopathic pulmonary fibrosis. The L-PF questionnaire was completed at baseline and week 52. The responsiveness of the Dyspnoea and Cough scores was evaluated by comparing changes in these scores with 52-week changes in three anchors: forced vital capacity % predicted and two self-reported items, one for global physical health and one for global quality of life. We used a triangulation approach including anchor-based and distribution-based methods to estimate meaningful change thresholds.

**Results** The analyses included 542 subjects with an L-PF Dyspnoea score at baseline and week 52, and 538 subjects with an L-PF Cough score at baseline and week 52. The L-PF Dyspnoea and Cough scores were responsive to change over 52 weeks. Triangulation of anchor-based and distribution-based estimates resulted in meaningful change thresholds of 6 to 7 points for the L-PF Dyspnoea score and 4 to 5 points for the L-PF Cough score to differentiate subjects who were stable or improved from those who deteriorated.

**Conclusion** These analyses support the responsiveness, one aspect of validity, of the L-PF Dyspnoea and Cough domains scores as measures of symptom severity in patients with progressive fibrosing ILDs. Estimates for meaningful change thresholds in these domain scores may be of value in interpreting the effects of interventions in these patients.

**Trial registration number** NCT02999178.

## INTRODUCTION

Idiopathic pulmonary fibrosis (IPF) is an inexorably progressive fibrosing interstitial

## Key messages

### What is already known on this topic

- The Living with Pulmonary Fibrosis (L-PF) questionnaire was developed to assess quality of life in patients with fibrosing interstitial lung diseases (ILDs), but little was known about its responsiveness to change.

### What this study adds

- In patients with progressive fibrosing ILDs, the L-PF Dyspnoea and Cough scores are responsive to changes in patients' health status and quality of life.

### How this study might affect research, practice or policy

- These analyses provide estimates for meaningful change thresholds for the L-PF Dyspnoea and Cough scores that may be of value in interpreting the effects of interventions.

lung disease (ILD).<sup>1</sup> A number of other ILDs may also be associated with a progressive fibrosing phenotype, characterised by an increasing extent of fibrosis, decline in lung function, worsening symptoms and quality of life and early mortality.<sup>2-4</sup> In patients with fibrosing ILDs, dyspnoea, cough and fatigue can affect patients' physical and emotional well-being and health-related quality of life (HRQL),<sup>5</sup> which tends to decline as patients' lung function worsens.<sup>6,7</sup>

Patient-centred outcomes are important tools for assessing the effects of disease and interventions on aspects of patients' lives, including symptoms and HRQL.<sup>8</sup> The Living with Idiopathic Pulmonary Fibrosis (L-IPF) questionnaire, which includes two modules that assess symptoms or their impacts, was



developed to assess health status and quality of life in patients with IPF. This questionnaire demonstrated sound psychometric properties in these patients, including discrimination between those with different disease severities.<sup>9</sup>

The Living with Pulmonary Fibrosis (L-PF) questionnaire is a slightly modified version of the L-IPF questionnaire. Like L-IPF, it has two modules, Symptoms and Impacts; the Symptoms module comprises three domains: Dyspnoea, Cough and Fatigue. L-PF is intended to be used in patients with all forms of progressive fibrosing ILD, including IPF; thus, the goal is for this questionnaire to replace the L-IPF questionnaire. Debriefing interviews of patients with progressive fibrosing ILDs other than IPF indicate that the L-PF has excellent face validity, that its concepts are relevant and that its items are understood as intended.<sup>10</sup>

The responsiveness of a patient-reported outcome, that is, its capacity to detect change in the target construct, as determined by clinically relevant outcomes or patients' perceptions (anchors) is an important aspect of its validation.<sup>11-13</sup> In this study, we evaluated the responsiveness of the L-PF Dyspnoea and Cough domain scores and estimated meaningful change thresholds (minimal clinically important differences) in these scores in patients with progressive fibrosing ILDs other than IPF.

**METHODS**  
**Trial design**

The INBUILD trial enrolled subjects with progressive fibrosing ILDs other than IPF. The trial design has been described and the protocol is publicly available.<sup>14</sup> Briefly, eligible subjects had a physician-diagnosed ILD other than IPF; reticular abnormality with traction bronchiectasis, with or without honeycombing, of >10% extent on high-resolution CT (HRCT); forced vital capacity (FVC) ≥45% predicted and diffusing capacity of the lungs for carbon monoxide ≥30% to <80% predicted. Subjects met one of the following criteria for ILD progression within the 24 months before screening despite management deemed appropriate in clinical practice: relative decline in FVC ≥10% predicted; relative decline in FVC ≥5 to <10% predicted and worsened respiratory symptoms and/or increased extent of fibrosis on HRCT; worsened respiratory symptoms and increased extent of fibrosis on HRCT. Subjects were randomised to receive nintedanib or placebo. The primary end point (annual rate of decline in FVC) was assessed over 52 weeks.

The L-PF questionnaire was completed at baseline and week 52. The L-PF questionnaire comprises 44 items: 23 in the Symptoms module and 21 in the Impacts module. Recall for items in the Symptoms module is the past 24 hours. Recall for items in the Impacts module is the past week. Domain and total scores range from 0 to 100, with higher scores indicating greater impairment. The L-PF questionnaire is accessible via: <https://eprovide.mapi-trust.org/instruments/>

**Table 1** Changes in L-PF Dyspnoea and Cough scores across strata of 52-week change in FVC % predicted strata

	Change in FVC % predicted†												F-statistic‡	Pairwise comparisons§									
	Large deterioration			Moderate deterioration			Minimal deterioration			Stable					Minimal improvement			Moderate improvement			Large improvement		
	N	Mean (SD) change from baseline	Mean (SD) change from baseline	N	Mean (SD) change from baseline	Mean (SD) change from baseline	N	Mean (SD) change from baseline	Mean (SD) change from baseline	N	Mean (SD) change from baseline	Mean (SD) change from baseline			N	Mean (SD) change from baseline	Mean (SD) change from baseline	N	Mean (SD) change from baseline	Mean (SD) change from baseline	N	Mean (SD) change from baseline	Mean (SD) change from baseline
L-PF Dyspnoea score	170	13.4 (17.2)	103 4.8 (15.5)	67	1.1 (12.9)	96	1.0 (13.7)	36	0.8 (9.0)	33	-4.0 (8.6)	25	-0.9 (14.4)	14.4***	1** , 2*** , 3*** , 4** , 5*** , 6**								
L-PF Cough score	168	8.8 (23.3)	102 0.2 (21.8)	67	-3.2 (20.1)	95	-2.3 (24.7)	36	-5.4 (30.4)	32	-6.5 (20.5)	26	-16.7 (20.5)	7.4***	2* , 3* , 6***								

L-PF Symptoms scores range from 0 to 100, with higher scores indicating greater impairment.

\*\*p<0.05, \*\*\*p<0.001.

†Large deterioration: decline >10% predicted; moderate deterioration: decline >5% and ≤10% predicted; minimal deterioration: decline >2% and ≤5% predicted; stable: decline ≤2% predicted or increase ≤2% predicted; minimal improvement: increase >2% and ≤5% predicted; moderate improvement: increase >5 and ≤10% predicted; large improvement: increase >10% predicted.

‡One-way analyses of variance (Generalised linear model) overall F-statistic.

§Pairwise comparisons using Scheffe's method: 1=large deterioration versus moderate deterioration, 2=large deterioration versus minimal deterioration, 3=large deterioration versus stable, 4=large deterioration versus minimal improvement, 5=large deterioration versus moderate improvement, 6=large deterioration versus large improvement, 7=moderate deterioration versus minimal deterioration, 8=moderate deterioration versus stable, 9=moderate deterioration versus minimal improvement, 10=moderate deterioration versus moderate improvement, 11=moderate deterioration versus large improvement, 12=minimal deterioration versus stable, 13=minimal deterioration versus minimal improvement, 14=minimal deterioration versus moderate improvement, 15=minimal deterioration versus large improvement, 16=stable deterioration versus minimal improvement, 17=stable deterioration versus moderate improvement, 18=stable deterioration versus large improvement, 19=minimal improvement versus moderate improvement, 20=minimal improvement versus large improvement, 21=moderate improvement versus large improvement.

FVC, forced vital capacity; L-PF, Living with Pulmonary Fibrosis.

living-with-pulmonary-fibrosis-l-pf-impacts-questionnaire and <https://eprovidemapi-trustorg/instruments/living-with-pulmonary-fibrosis-l-pf-symptoms-questionnaire>.

## Analyses

Analyses were conducted in subjects who received  $\geq 1$  dose of trial medication and had an L-PF Dyspnoea domain score (for analyses of this score) or L-PF Cough domain score (for analyses of this score) at baseline and at week 52. Data from the nintedanib and placebo groups were pooled. The responsiveness of the L-PF Dyspnoea and Cough scores was evaluated by comparing mean changes at week 52 across changes in three anchors at week 52: (1) absolute change from baseline in FVC % predicted; (2) absolute change from baseline in global physical health self-assessment (L-PF Impacts module, item 20: on average, over the last 7 days, how have you felt in terms of physical health? Scale: 0 (extremely poor) to 4 (excellent)), and 3) absolute change from baseline in the global quality of life self-assessment (L-PF Impacts module, item 21: on average, over the last 7 days, how has your quality of life been? Scale: 0 (extremely poor) to 4 (excellent)).

Changes in FVC % predicted at week 52 were categorised as follows: large deterioration (decline  $>10\%$  predicted); moderate deterioration (decline  $>5\%$  and  $\leq 10\%$  predicted); minimal deterioration (decline  $>2\%$  and  $\leq 5\%$  predicted); stable (decline  $\leq 2\%$  predicted or increase  $\leq 2\%$  predicted); minimal improvement (increase  $>2\%$  and  $\leq 5\%$  predicted); moderate improvement (increase  $>5\%$  and  $\leq 10\%$  predicted); large improvement (increase  $>10\%$  predicted). Changes in each global rating item at week 52 ranged from  $-4$  to  $+4$  and were categorised as follows: large deterioration ( $-3$  or  $-4$ ); moderate deterioration ( $-2$ ); minimal deterioration ( $-1$ ); stable ( $0$ ); minimal improvement ( $+1$ ); moderate improvement ( $+2$ ); large improvement ( $+3$ ,  $+4$ ). We used one-way analysis of variance models to examine whether changes in the Dyspnoea and Cough domain scores differed significantly across anchor strata. Scheffe's method was used for pairwise comparisons.

For the threshold analysis, point estimates for meaningful change were considered to be half-way between the mean changes in scores in subjects who were stable and in subjects who had minimal decline in FVC % predicted or minimal/moderate deterioration in the global rating anchors. To refine the thresholds of meaningful change for the global rating anchors, we also considered the half-way point between the mean changes in scores in the stable and minimal deterioration groups.

We used receiver operating characteristic (ROC) curves and Youden's index<sup>15</sup> to identify thresholds that maximised sensitivity and specificity of the Dyspnoea or Cough scores to differentiate subjects who deteriorated (decline in FVC  $>2\%$  predicted or change in global rating anchors of  $-4$  to  $-1$ ) from those who were

**Table 2** Changes in L-PF Dyspnoea and Cough scores across strata of 52-week change in global physical health score

	Change in L-PF global physical health score†										F-statistic‡	Pairwise comparisons§
	Large deterioration		Minimal/Moderate deterioration		Stable		Minimal/Moderate improvement		Large improvement			
	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline		
L-PF Dyspnoea score	6	17.6 (28.8)	159	8.7 (16.8)	244	5.7 (14.5)	120	0.6 (15.6)	8	1.0 (8.6)	5.7***	6**
L-PF Cough score	5	5.8 (27.7)	157	4.8 (24.8)	242	1.5 (23.3)	121	-5.9 (23.8)	8	0.0 (22.6)	3.6**	6**

L-PF Symptoms scores range from 0 to 100, with higher scores indicating greater impairment.

\*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

†Question 20: on average, over the last 7 days, how have you felt in terms of physical health? Scale: 0 (extremely poor) to 4 (excellent). Large deterioration: change of  $-3$  or  $-4$ ; minimal/moderate deterioration: change of  $-2$  or  $-1$ ; stable: change of  $0$ ; minimal/moderate improvement: change of  $+1$  or  $+2$ ; large improvement: change of  $+3$  or  $+4$ .

‡One-way analyses of variance (generalised linear model) overall F-statistic.

§Pairwise comparisons using Scheffe's method: 1=large deterioration versus minimal/moderate deterioration, 2=large deterioration versus stable, 3=large deterioration versus minimal/moderate improvement, 4=large deterioration versus large improvement, 5=minimal/moderate deterioration versus stable, 6=minimal/moderate deterioration versus minimal/moderate improvement, 7=minimal/moderate deterioration versus large improvement, 8=stable versus minimal/moderate improvement, 9=stable versus large improvement, 10=minimal/moderate improvement versus large improvement.

L-PF, Living with Pulmonary Fibrosis.

stable/improved (improvement or decline in FVC  $\leq 2\%$  predicted or change in global rating anchors of 0 to +4).

Distribution-based analyses were performed to provide supplementary results. We evaluated the SEM, estimated as the baseline SD of the measure multiplied by the square root of 1 minus its reliability coefficient, and  $0.2 \times \text{SD}$  and  $0.5 \times \text{SD}$  of the scores at baseline. One SEM may be considered a meaningful change threshold<sup>16 17</sup> and changes of  $0.5 \times \text{SD}$  and  $0.2 \times \text{SD}$  may be considered upper and lower boundaries for a meaningful change.<sup>18</sup>

### Patient and public involvement

Patients and/or the public were not involved in the design, conduct, reporting or dissemination plans of this research.

### RESULTS

A total of 663 subjects were enrolled in the INBUILD trial at 153 sites in 15 countries. Their baseline characteristics have been described.<sup>14</sup> Briefly, mean (SD) age was 65.8 (9.8) years and FVC was 69.0 (15.6) % predicted; 53.7% of subjects were male and 62.1% had a usual interstitial pneumonia-like fibrotic pattern on HRCT. The most common diagnoses were hypersensitivity pneumonitis (26.1%), autoimmune disease-related ILDs (25.6%), idiopathic non-specific interstitial pneumonia (18.9%) and unclassifiable ILD (17.2%). A total of 542 subjects with an L-PF Dyspnoea score at baseline and week 52, and 538 subjects with an L-PF Cough score at baseline and week 52, were included in these analyses.

### Responsiveness of L-PF Dyspnoea and Cough scores

There were large and statistically significant differences in changes in Dyspnoea and Cough scores between subjects with stable versus large deterioration in FVC % predicted and between subjects with minimal versus large deterioration in FVC % predicted (table 1). There was a statistically significant difference in change in Dyspnoea score between subjects with moderate versus large deterioration in FVC % predicted. There were no statistically significant differences in changes in Dyspnoea or Cough scores between subjects with stable versus minimal deterioration in FVC % predicted. Changes in Dyspnoea and Cough scores were significantly different between subjects with minimal/moderate deterioration versus minimal/moderate improvement in either global rating anchor (tables 2 and 3).

### Meaningful change thresholds in L-PF Dyspnoea score

For the Dyspnoea domain, the half-way points between changes in scores for subjects who were stable and those with minimal deterioration in FVC % predicted, global physical health score and global quality of life score were 1.1, 7.1 and 7.6, respectively. Similar half-way points were observed for the global rating anchors when minimal/

**Table 3** Changes in L-PF Dyspnoea and Cough scores across strata of 52-week change in global quality of life score

	Change in L-PF global quality of life score†										
	Large deterioration		Minimal/Moderate deterioration		Stable		Minimal/Moderate improvement		Large improvement		Pairwise comparisons‡
	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	N	Mean (SD) change from baseline	
L-PF Dyspnoea score	7	10.8 (29.6)	167	9.4 (16.1)	240	5.4 (14.4)	115	-0.1 (16.1)	7	4.1 (15.7)	6***, 8*
L-PF Cough score	7	1.8 (31.8)	165	6.9 (24.4)	238	-0.5 (21.2)	115	-4.9 (27.2)	7	-6.0 (25.6)	6**

L-PF Symptoms scores range from 0 to 100, with higher scores indicating greater impairment.

\* $P < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

†Question 21: on average, over the last 7 days, how has your quality of life been? Scale: 0 (extremely poor) to 4 (excellent). Large deterioration: change of -3 or -4; minimal/moderate deterioration: change of -2 or -1; stable: change of 0; minimal/moderate improvement: change of +1 or +2; large improvement: change of +3 or +4.

‡One-way analyses of variance (generalised linear model) overall F-statistic.

§Pairwise comparisons using Scheffe's method: 1=large deterioration versus minimal/moderate deterioration, 2=large deterioration versus stable, 3=large deterioration versus minimal/moderate improvement, 4=large deterioration versus large improvement, 5=minimal/moderate deterioration versus stable, 6=minimal/moderate deterioration versus minimal/moderate improvement, 7=minimal/moderate deterioration versus large improvement, 8=stable versus minimal/moderate improvement, 9=stable versus large improvement, 10=minimal/moderate improvement versus large improvement.

L-PF, Living with Pulmonary Fibrosis.

**Table 4** Meaningful change thresholds for L-PF Dyspnoea and Cough domain scores

	L-PF Dyspnoea score			L-PF Cough score		
	Change in FVC % predicted	Change in L-PF global physical health score	Change in L-PF global quality of life score	Change in FVC % predicted	Change in L-PF global physical health score	Change in L-PF global quality of life score
<b>Patients with minimal deterioration versus patients who were stable</b>						
Minimal deterioration	1.1 (12.9)	8.5 (17.2)	9.8 (15.7)	-3.2 (20.1)	4.1 (22.9)	7.5 (23.0)
Stable	1.0 (13.7)	5.7 (14.5)	5.4 (14.4)	-2.3 (24.7)	1.5 (23.3)	-0.5 (21.2)
<b>Half-way point</b>	<b>1.1</b>	<b>7.1</b>	<b>7.6</b>	<b>-2.8</b>	<b>2.8</b>	<b>3.5</b>
<b>Patients with minimal/moderate deterioration versus patients who were stable</b>						
Minimal/Moderate deterioration	n/a	8.7 (16.8)	9.4 (16.1)	n/a	4.8 (24.8)	6.9 (24.4)
Stable	n/a	5.7 (14.5)	5.4 (14.4)	n/a	1.5 (23.3)	-0.5 (21.2)
<b>Half-way point</b>		<b>7.2</b>	<b>7.4</b>		<b>3.2</b>	<b>3.2</b>

Mean (SD) changes from baseline.

FVC, forced vital capacity; L-PF, Living with Pulmonary Fibrosis; n/a, not analysed.

moderate deterioration was considered instead of minimal deterioration (table 4).

For the Dyspnoea score, ROC analyses revealed meaningful change thresholds between deterioration and stability/improvement of 5.6 for FVC % predicted, 6.3 for global physical health and 1.7 for global quality of life (table 5).

In distribution-based estimates of thresholds of meaningful change in the Dyspnoea score, the SEM was 4.4,  $0.2 \times SD$  was 4.3 and  $0.5 \times SD$  was 10.8.

Triangulation of the anchor-based and distribution-based estimates for the Dyspnoea domain score resulted in a meaningful change threshold of 6 to 7 points to

differentiate subjects who were stable from those who deteriorated.

#### Meaningful change thresholds in L-PF Cough score

For the Cough domain, the half-way points between changes in scores for subjects who were stable and those with minimal deterioration in FVC % predicted, global physical health score and global quality of life score were -2.8, 2.8 and 3.5, respectively (table 4). Similar half-way points were observed for the global ratings anchors when minimal/moderate deterioration was considered instead of minimal deterioration (table 4).

**Table 5** Results from receiver operating characteristic (ROC) curve analysis: sensitivity and specificity of L-PF questionnaire Dyspnoea and Cough scores to distinguish deterioration (vs stability/improvement) based on FVC % predicted, L-PF global physical health and quality of life scores

	Youden's index*	Cut-point	Sensitivity	Specificity
<b>Deterioration in FVC % predicted</b>				
L-PF Dyspnoea score	0.26	5.6	0.49	0.77
L-PF Cough score	0.18	4.2	0.50	0.67
<b>Deterioration in L-PF questionnaire global physical health score</b>				
L-PF Dyspnoea score	0.19	6.3	0.52	0.67
L-PF Cough score	0.12	16.7	0.35	0.78
<b>Deterioration in L-PF questionnaire global quality of life score</b>				
L-PF Dyspnoea score	0.18	1.7	0.67	0.51
L-PF Cough score	0.16	4.2	0.55	0.61

\*Index defining the cut-point that maximises sensitivity and specificity.

FVC, forced vital capacity; L-PF, Living with Pulmonary Fibrosis.



For the Cough score, ROC analyses revealed meaningful change thresholds between deterioration and stability/improvement of 4.2 for FVC % predicted, 16.7 for global physical health and 4.2 for global quality (table 5).

In distribution-based estimates of thresholds of meaningful change in the Cough score, SEM was 8.6,  $0.2 \times \text{SD}$  was 5.3 and  $0.5 \times \text{SD}$  was 13.3.

Triangulation of the anchor-based and distribution-based estimates for the Cough domain score resulted in a meaningful change threshold of 4 to 5 points to differentiate subjects who were stable from those who deteriorated.

## DISCUSSION

Our analyses suggest that in patients with progressive fibrosing ILDs other than IPF, the Dyspnoea and Cough domain scores from the L-PF questionnaire Symptoms module are responsive to changes in disease severity and in patients' perceptions of their physical health and quality of life. We observed significant differences in changes in L-PF Dyspnoea and Cough scores between subjects who had a large deterioration in FVC % predicted versus those with stable FVC % predicted, and between subjects who experienced deterioration versus improvement in global assessment anchors.

There is no consensus on the best approach to estimating meaningful change thresholds for patient-reported outcomes.<sup>13 18</sup> Food and Drug Administration guidance recommends that anchor-based approaches incorporate 'patient ratings' of change<sup>12</sup>; however, such transition items, which require patients to assess their current state, recall their prior state and mentally subtract the difference (eg, "Is your shortness of breath a lot better/the same/a lot worse?"), are fraught with problems. Ideally, the correlation between the transition item and baseline score is equal and opposite to the correlation between the transition item and the score at follow-up, but with recall periods of longer than 4 weeks, transition ratings tend to be (inappropriately) highly correlated with the patient's current state.<sup>19</sup> The two patient response anchors we used alleviated this potential for bias by asking patients to rate their state at baseline and at week 52; we then performed the subtraction to yield the transition item.

For many transition items, stability and degree of change are arbitrarily defined by the investigator. Some investigators may consider 'somewhat worse/better' to represent a minimal change, while others may consider 'a bit worse/better' or 'minimally worse/better' to be a minimal change. How patients interpret such descriptors, and how investigators categorise anchors, can affect estimates of meaningful change thresholds. For example, when using a 15-point quality of life transition item with ratings ranging from -7 to +7, ratings of -1 to +1 have been considered to represent no change and ratings of -3 to -2, +2 and +3 to represent minimally important

changes,<sup>20 21</sup> but meaningful change estimates may have been different if stability had been defined as a rating of 0 and minimally important changes as ratings of -2 to -1, +1 and +2. For our global rating anchors, we considered a change of 0 to represent stability and changes of -1 to -2, +1 and +2 to represent minimal/moderate change. Some patients with transition scores of 0 may have changed minimally and some with transition scores of 1 or 2 may have been stable. We attempted to account for this inherent uncertainty by using a half-way point approach rather than simply subtracting mean scores between groups of interest.

As patients with progressive ILDs are unlikely to experience improvement in disease status, in the ROC analyses, we identified a change threshold between worsening and stability/improvement. This approach aligns with the clinical behaviour of progressive ILDs and with current therapeutic approaches, which slow rather than reverse disease progression.

Change in FVC is used as a primary end point in clinical trials to assess the efficacy of treatments for ILDs.<sup>14 22-25</sup> A decline in FVC is associated with mortality.<sup>2 26-28</sup> While there is no established definition of ILD progression, absolute declines of >5% or >10% in FVC % predicted are widely regarded as indicating progression,<sup>26 28-30</sup> although smaller declines may also be relevant. Scores from patient-reported outcomes that assess symptoms or HRQL typically correlate weakly with FVC in patients with ILDs,<sup>31-33</sup> suggesting that these measures yield information unique from physiological measures of ILD severity. This suggests that although commonly used as an anchor in validation studies, FVC may not be a suitable anchor in all circumstances.

Strengths of our analyses include the use of a large and heterogeneous population of subjects with progressive fibrosing ILDs. The use of triangulation that incorporated both anchor-based and distribution-based approaches aligns with accepted methodology, including from regulatory bodies, but we acknowledge that distribution-based methods may overestimate meaningful change thresholds.<sup>34</sup> Limitations include that the trial was not designed to evaluate the measurement properties of patient-reported outcomes, so additional metrics that could have been used as anchors were not included. For example, another cough-specific patient-reported outcome would have been a more appropriate anchor for the Cough domain. The content validity of the L-PF questionnaire has not been demonstrated for all the languages and cultures that participated in the trial. Whether our findings are applicable to patients with fibrosing ILDs beyond those who met the inclusion criteria for the INBUILD trial is unknown.

In conclusion, our analyses support the responsiveness of the Dyspnoea and Cough domains of the L-PF questionnaire Symptoms module as measures of symptom severity in patients with progressive fibrosing ILDs. Estimates of meaningful change thresholds in these scores may be of value in interpreting the effects of interventions in these

patients. Additional analyses are encouraged to confirm or refine these findings.

#### Author affiliations

<sup>1</sup>National Jewish Health, Denver, Colorado, USA

<sup>2</sup>Evidera (Pharmaceutical Product Development, LLC), Bethesda, Maryland, USA

<sup>3</sup>Boehringer Ingelheim International GmbH, Ingelheim am Rhein, Germany

<sup>4</sup>Boehringer Ingelheim Pharma GmbH & Co. KG, Biberach, Germany

<sup>5</sup>National Hospital Organization Kinki-Chuo Chest Medical Center, Osaka, Japan

**Acknowledgements** Writing assistance was provided by Elizabeth Ng and Wendy Morris of FleishmanHillard, London, UK, which was contracted and funded by BI. BI was given the opportunity to review the manuscript for medical and scientific accuracy as well as intellectual property considerations. JJS was at National Jewish Health during the writing of this manuscript and now an employee of Boehringer-Ingelheim.

**Contributors** JJS, DMB, KR and HM contributed to the planning of the analyses. DMB and HM contributed to the conduct of the analyses. JJS, DMB, KR, HM, MB and YI contributed to the interpretation of the data and the writing of the manuscript. HM is responsible for the overall content of the manuscript as guarantor.

**Funding** The INBUILD trial was funded by Boehringer Ingelheim International GmbH (BI). The authors meet criteria for authorship as recommended by the International Committee of Medical Journal Editors (ICMJE). The authors did not receive payment for development of this manuscript.

**Competing interests** Yes.

**Patient consent for publication** Not applicable.

**Ethics approval** The trial was carried out in compliance with the protocol, the principles of the Declaration of Helsinki and the Harmonised Tripartite Guideline for Good Clinical Practice of the International Conference on Harmonisation. The protocol was approved by an independent ethics committee or institutional review board at each participating centre (additional details can be found in online supplemental appendix 1). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request. To ensure independent interpretation of clinical study results, Boehringer Ingelheim grants all external authors access to relevant material, including participant-level clinical study data, as needed by them to fulfil their role and obligations as authors under the ICMJE criteria. Clinical study documents and participant clinical study data are available to be shared on request after publication of the primary manuscript in a peer-reviewed journal, and if regulatory activities are complete and other criteria met as per the BI Policy on Transparency and Publication of Clinical Study Data (<https://www.mystudywindow.com/msw/datasharing>). Bona fide, qualified scientific and medical researchers are eligible to request access to the clinical study data with corresponding documentation describing the structure and content of the datasets. On approval, and governed by a Legal Agreement, data are shared in a secured data-access system for a period of 1 year, which may be extended on request. Prior to providing access, clinical study documents and data will be examined, and, if necessary, redacted and de-identified, to protect the personal data of study participants and personnel, and to respect the boundaries of informed consent. Researchers should use the <https://vivli.org/> link to request access to study data and visit <https://www.mystudywindow.com/msw/datasharing> for further information.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/4.0/>.

#### ORCID iDs

Jeffrey J Swigris <http://orcid.org/0000-0002-2643-8110>

Yoshikazu Inoue <http://orcid.org/0000-0003-3994-874X>

#### REFERENCES

- Raghu G, Remy-Jardin M, Myers JL, *et al*. Diagnosis of idiopathic pulmonary fibrosis. An official ATS/ERS/JRS/ALAT clinical practice guideline. *Am J Respir Crit Care Med* 2018;198:e44–68.
- Brown KK, Martinez FJ, Walsh SLF, *et al*. The natural history of progressive fibrosing interstitial lung diseases. *Eur Respir J* 2020;55:2000085.
- Faverio P, Piluso M, De Giacomo F, *et al*. Progressive fibrosing interstitial lung diseases: prevalence and characterization in two Italian referral centers. *Respiration* 2020;99:1–8.
- Nasser M, Larrieu S, Si-Mohamed S, *et al*. Progressive fibrosing interstitial lung disease: a clinical cohort (the progress study). *Eur Respir J* 2021;57:2002718.
- Swigris JJ, Brown KK, Abdulqawi R, *et al*. Patients' perceptions and patient-reported outcomes in progressive-fibrosing interstitial lung diseases. *Eur Respir Rev* 2018;27:180075.
- Kreuter M, Wuys WA, Wijsenbeek M, *et al*. Health-related quality of life and symptoms in patients with IPF treated with nintedanib: analyses of patient-reported outcomes from the INPULSIS® trials. *Respir Res* 2020;21:36.
- Maqhuzu PN, Szentes BL, Kreuter M, *et al*. Determinants of health-related quality of life decline in interstitial lung disease. *Health Qual Life Outcomes* 2020;18:334.
- Aronson KI, Danoff SK, Russell A-M, *et al*. Patient-centered outcomes research in interstitial lung disease: an official American thoracic Society research statement. *Am J Respir Crit Care Med* 2021;204:e3–23.
- Swigris JJ, Andrae DA, Churney T, *et al*. Development and initial validation analyses of the living with idiopathic pulmonary fibrosis questionnaire. *Am J Respir Crit Care Med* 2020;202:1689–97.
- Swigris J, Cutts K, Male N, *et al*. The living with pulmonary fibrosis questionnaire in progressive fibrosing interstitial lung disease. *ERJ Open Res* 2021;7:00145–2020.
- Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003;56:395–407.
- Food and Drug Administration. Guidance for industry. patient-reported outcome measures: use in medical product development to support labeling claims, 2009. Available: <https://www.fda.gov/media/77832/download>
- Kalluri M, Luppi F, Vancheri A, *et al*. Patient-reported outcomes and patient-reported outcome measures in interstitial lung disease: where to go from here? *Eur Respir Rev* 2021;30:210026.
- Flaherty KR, Wells AU, Cottin V, *et al*. Nintedanib in progressive fibrosing interstitial lung diseases. *N Engl J Med* 2019;381:1718–27.
- Fluss R, Faraggi D, Reiser B. Estimation of the Youden index and its associated cutoff point. *Biom J* 2005;47:458–72.
- Wyrwich KW, Nienaber NA, Tierney WM, *et al*. Linking clinical relevance and statistical significance in evaluating intra-individual changes in health-related quality of life. *Med Care* 1999;37:469–78.
- Wyrwich KW, Tierney WM, Wolinsky FD. Further evidence supporting an SEM-based criterion for identifying meaningful intra-individual changes in health-related quality of life. *J Clin Epidemiol* 1999;52:861–73.
- Revicki D, Hays RD, Cella D, *et al*. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008;61:102–9.
- Guyatt GH, Norman GR, Juniper EF, *et al*. A critical look at transition ratings. *J Clin Epidemiol* 2002;55:900–8.
- Juniper EF, Guyatt GH, Griffith LE, *et al*. Interpretation of rhinoconjunctivitis quality of life questionnaire data. *J Allergy Clin Immunol* 1996;98:843–5.
- Juniper EF, Guyatt GH, Willan A, *et al*. Determining a minimal important change in a disease-specific quality of life questionnaire. *J Clin Epidemiol* 1994;47:81–7.
- Richeldi L, du Bois RM, Raghu G, *et al*. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2071–82.
- King TE, Bradford WZ, Castro-Bernardini S, *et al*. A phase 3 trial of pirfenidone in patients with idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370:2083–92.
- Distler O, Highland KB, Gahlemann M, *et al*. Nintedanib for systemic sclerosis-associated interstitial lung disease. *N Engl J Med* 2019;380:2518–28.



- 25 Maher TM, Corte TJ, Fischer A, *et al.* Pirfenidone in patients with unclassifiable progressive fibrosing interstitial lung disease: a double-blind, randomised, placebo-controlled, phase 2 trial. *Lancet Respir Med* 2020;8:147–57.
- 26 Solomon JJ, Chung JH, Cosgrove GP, *et al.* Predictors of mortality in rheumatoid arthritis-associated interstitial lung disease. *Eur Respir J* 2016;47:588–96.
- 27 Goh NS, Hoyles RK, Denton CP, *et al.* Short-Term pulmonary function trends are predictive of mortality in interstitial lung disease associated with systemic sclerosis. *Arthritis Rheumatol* 2017;69:1670–8.
- 28 Gimenez A, Storrer K, Kuranishi L, *et al.* Change in FVC and survival in chronic fibrotic hypersensitivity pneumonitis. *Thorax* 2018;73:391–2.
- 29 Khanna D, Mittoo S, Aggarwal R, *et al.* Connective tissue disease-associated Interstitial Lung Diseases (CTD-ILD) - report from OMERACT CTD-ILD Working Group. *J Rheumatol* 2015;42:2168–71.
- 30 Paterniti MO, Bi Y, Rekić D, *et al.* Acute exacerbation and decline in forced vital capacity are associated with increased mortality in idiopathic pulmonary fibrosis. *Ann Am Thorac Soc* 2017;14:1395–402.
- 31 Swigris JJ, Wilson H, Esser D, *et al.* Psychometric properties of the St George's respiratory questionnaire in patients with idiopathic pulmonary fibrosis: insights from the INPULSIS trials. *BMJ Open Respir Res* 2018;5:e000278.
- 32 Prior TS, Hoyer N, Hilberg O, *et al.* Responsiveness and minimal clinically important difference of SGRQ-I and K-BILD in idiopathic pulmonary fibrosis. *Respir Res* 2020;21:91.
- 33 Kim JW, Clark A, Biring SS, *et al.* Psychometric properties of patient reported outcome measures in idiopathic pulmonary fibrosis. *Chron Respir Dis* 2021;18:14799731211033925.
- 34 Stratford PW, Riddle DL. When minimal detectable change exceeds a diagnostic test-based threshold change value for an outcome measure: resolving the conflict. *Phys Ther* 2012;92:1338–47.