Psychometric properties of the St George's Respiratory Questionnaire in patients with idiopathic pulmonary fibrosis

Jeffrey J. Swigris[1], Dirk Esser[2], Hilary Wilson[3], Craig S. Conoscenti[4], Hendrik Schmidt[5], Wibke Stansen[5], Nancy Kline Leidy[3] and Kevin K. Brown[1]

Affiliations: [1]National Jewish Health, Denver, CO, USA. [2]Boehringer Ingelheim GmbH, Ingelheim am Rhein, Germany. [3]Evidera, Bethesda, MD, USA. [4]Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA. [5]Boehringer Ingelheim GmbH & Co. KG, Ingelheim am Rhein, Germany.

Correspondence: Jeffrey J. Swigris, Autoimmune Lung Center and Interstitial Lung Disease Program, National Jewish Health, 1400 Jackson Street, Denver, CO 80206, USA. E-mail: swigrisj@njc.org

ABSTRACT The St George's Respiratory Questionnaire (SGRQ) has been used to measure health-related quality of life (HRQoL) in patients with idiopathic pulmonary fibrosis (IPF).

This analysis evaluated the psychometric properties of the SGRQ using data from 428 patients with IPF who participated in a 12-month, randomised, placebo-controlled phase II trial of nintedanib.

Internal consistency (Cronbach's α) was 0.91 for SGRQ total and >0.70 for domain scores. Test–retest reliability (intraclass correlation coefficients) was 0.77, 0.77, 0.69 and 0.66 for SGRQ total, activity, impact and symptoms scores, respectively. Construct validity of SGRQ total and domain scores was supported by weak to moderate cross-sectional correlations with the Medical Research Council dyspnoea scale (0.32–0.55), 6-min walk test distance (−0.25–−0.34), percentage predicted forced vital capacity (−0.11–−0.15) and measures of gas exchange (−0.26–0.03). There was some evidence that the SGRQ total score was sensitive to detecting change.

The reliability, construct validity and responsiveness of the SGRQ in patients with IPF suggest that this is an acceptable measure of HRQoL in patients with IPF.

Introduction

Idiopathic pulmonary fibrosis (IPF) is a fibrosing interstitial pneumonia characterised by progressively worsening dyspnoea and lung function [1]. In the USA, the annual incidence of IPF has been estimated as 6.8–8.8 cases per 100 000 population using narrow case definitions and as 16.3–17.4 cases per 100 000 population using broad case definitions [2]. Although IPF has a poor prognosis, with a median survival time from diagnosis of 2–3 years, the clinical course of IPF is extremely variable [1, 3]. Symptoms interfere with daily activities, resulting in a loss of independence and inducing fears about the future, thus impairing patients' health-related quality of life (HRQoL) [4, 5].

IPF is a progressive disease with no cure; as the disease worsens, patients are forced to readjust their lives to their declining functional capacity, and this makes HRQoL an important end-point to target in the clinical and research arenas [6]. The St George's Respiratory Questionnaire (SGRQ) was originally developed for use in patients with chronic obstructive pulmonary disease and asthma [7], but it has been used to assess HRQoL in patients with IPF [8]. The SGRQ is a 50-item questionnaire assessing three domains of HRQoL: symptoms (frequency and severity of respiratory symptoms), activity (effects of breathlessness on physical activities and *vice versa*) and impact (psychosocial impact of the disease). The SGRQ total score is a composite of the three domain scores. Scores are weighted and range from 0 to 100, with higher scores indicating a poorer HRQoL. Findings to date suggest that the SGRQ may be useful as a measure of HRQoL in clinical trials in patients with IPF assessing the effect of treatment on lung function decline and the preservation of HRQoL [8].

The aim of this analysis was to evaluate the psychometric properties of SGRQ scores in patients with IPF who participated in a 12-month, international phase II trial of a then investigational therapy: the TOMORROW trial of nintedanib [9].

Materials and methods

Patient population

Data from an international, randomised, double-blind, placebo-controlled phase II trial investigating the efficacy and safety of nintedanib (Boehringer Ingelheim, Ingelheim, Germany) as a treatment for IPF (the TOMORROW trial) were analysed. The design and efficacy and safety results from this trial, including the effect of nintedanib on SGRQ total and domain scores, have been published [9].

Eligibility criteria for enrolment in the TOMORROW trial included a diagnosis of IPF within 5 years prior to screening, a forced vital capacity (FVC) ⩾50% predicted, a diffusing capacity of the lung for carbon monoxide (D_{LCO}) 30–79% predicted and an arterial oxygen tension (P_{aO_2}) ⩾55 mmHg (sea level to 1500 m) or 50 mmHg (>1500 m) when breathing ambient air.

The trial was conducted in accordance with the principles of the Declaration of Helsinki and the Harmonised Tripartite Guideline for Good Clinical Practice from the International Conference on Harmonisation, and was approved by local authorities. Written informed consent was obtained from all participants. In total, 432 patients from 25 countries were randomised to receive one of four doses of nintedanib (50 mg once a day, 50 mg twice a day, 100 mg twice a day or 150 mg twice a day) or placebo for 12 months. Four patients were randomised but did not receive treatment. Data from all the treated patients were pooled (n=428) in the current analysis.

Outcome measures

We used data for FVC % pred, D_{LCO} % pred and resting peripheral arterial oxygen saturation measured by pulse oximetry (S_{pO_2}) obtained at baseline and weeks 24 and 52 for our analyses. We used data for P_{aO_2} and arterial carbon dioxide tension (P_{aCO_2}) obtained at baseline and week 52.

For the 6-min walk test, we analysed data obtained from assessments conducted at baseline and weeks 24 and 52. This test measures the distance an individual can walk in 6 min (6-min walk distance (6MWD)) and is considered a reliable measure of functional exercise capacity in patients with IPF [10]. The test was conducted according to American Thoracic Society criteria [11]; if a patient needed supplemental oxygen during the baseline test to avoid hypoxaemia, that flow rate was used for all subsequent tests.

Results of the SGRQ and the Medical Research Council (MRC) dyspnoea scale [12], completed at baseline and weeks 24 and 52, were analysed. The MRC dyspnoea scale comprises five statements that describe the range of respiratory disability from none (grade 1) to almost complete incapacity (grade 5). Both the SGRQ and the MRC dyspnoea scale were completed before any other trial-related procedures.

Statistical analysis

Psychometric properties, including internal consistency, test–retest reliability, construct validity, known-groups validity and responsiveness, were examined. All analyses were exploratory and performed *post hoc.*

Internal consistency

Cronbach's α coefficient was calculated to determine the internal consistency of SGRQ scores at baseline; values >0.7 are generally considered indicative of a homogeneous scale [13].

Test–retest reliability

The test–retest reliability of SGRQ scores was determined by calculating the intraclass correlation coefficient (ICC) and effect sizes in patients deemed clinically stable. Stable clinical status was defined as a change in FVC % pred of ⩽2% from baseline to week 52. We selected a threshold of 2% to define stable patients based on an estimated minimal clinically important difference for FVC % pred of 2–6% [14]. ICC values >0.7 are generally considered acceptable for establishing test–retest reliability [15]. We considered a small effect size (<0.2) to indicate stability of SGRQ scores between baseline and week 52.

Construct validity

Construct validity, both cross-sectional and longitudinal, was evaluated by examining the magnitude of correlations (Spearman coefficients) between SGRQ scores and the MRC dyspnoea score, 6MWD, FVC % pred and measures of gas exchange (D_{LCO} % pred, S_{pO_2}, P_{aO_2} and P_{aCO_2}). Cross-sectional assessments were performed at baseline, while longitudinal assessments were evaluated as change from baseline to week 24 or 52. We considered the strength of correlations per convention: weak <0.30, moderate 0.30–0.60 and strong >0.60 [16].

Known-groups validity

Although there are no established categories for disease severity in IPF, known-groups validity was assessed by examining differences in baseline mean SGRQ total scores between patients stratified by FVC % pred (⩽70% *versus* >70%), D_{LCO} % pred (⩽55% *versus* >55%) and Composite Physiologic Index (CPI) (⩽45 *versus* >45). Known-groups analyses were assessed by descriptive comparison of SGRQ total scores (box plots) at baseline.

Responsiveness

We used an anchor-based method to assess the responsiveness of the SGRQ. Patients were stratified by change in FVC % pred from baseline to week 52 (deterioration of >10%, deterioration of >5–⩽10%, deterioration of >2–⩽5%, deterioration or improvement of ⩽2% (stable), improvement of >2–⩽5%, improvement of >5–⩽10 and improvement of >10%) and by change in MRC dyspnoea score from baseline to week 52 (decrease, no change and increase). The ability to detect change was assessed using ANCOVA of changes in SGRQ total scores by changes in FVC % pred or MRC dyspnoea score categories at week 52.

Handling of missing data

If the entire questionnaire was missing at baseline, or if the only assessment was at baseline, then the patient was excluded from the analysis for all aspects of the questionnaire. Otherwise, "last observation carried forward" (LOCF) methodology was applied for the entire questionnaire. If an individual domain was missing at baseline, or if the only assessment was at baseline, then the patient was excluded from the analysis for all aspects of the questionnaire. If there were more than two, four or six missing items in the symptoms, activity or impact domains, respectively, then the domain was set to missing. Otherwise, LOCF methodology for the individual domain was applied.

Results

Patients

Patient demographics and baseline characteristics are summarised in online supplementary table S1.

Reliability

Internal consistency

Cronbach's α coefficients were ⩾0.83 for the SGRQ total, activity and impact scores, and 0.75 for the symptoms score (table 1).

TABLE 1 Internal consistency of St George's Respiratory Questionnaire (SGRQ) scores

SGRQ score	Cronbach's α (standardised)
Total	0.91
Activity domain	0.84
Impact domain	0.83
Symptoms domain	0.75

TABLE 2 Correlations (Spearman coefficients) between St George's Respiratory Questionnaire (SGRQ) scores and Medical Research Council (MRC) dyspnoea scores

SGRQ score	MRC dyspnoea score		
	Baseline	Change from baseline to week 24	Change from baseline to week 52
Total	0.55***	0.32***	0.35***
Activity domain	0.58***	0.34***	0.39***
Impact domain	0.50***	0.23***	0.26***
Symptoms domain	0.32***	0.22***	0.25***

***: $p < 0.001$.

Test–retest reliability

Test–retest reliability over a 52-week period was acceptable for the SGRQ total and activity scores (ICC 0.77 for both scores) in patients with stable disease. The ICCs for the SGRQ impact score (0.69) and symptoms score (0.66) were close to, but did not meet the threshold of >0.7. For patients with stable disease, the corresponding effect sizes for change in SGRQ scores from baseline to week 52 were small (−0.092, 0.148, −0.142 and −0.058 for the symptoms, activity, impact and total scores, respectively).

Validity
Construct validity

All cross-sectional correlations between SGRQ scores and the MRC dyspnoea score were moderate, in the expected direction and statistically significant (table 2). Correlations between changes in SGRQ total and activity scores and change in the MRC dyspnoea score were moderate, in the expected direction and statistically significant. Correlations between the changes in SGRQ impact and symptoms scores and change in the MRC dyspnoea score were weak, but statistically significant and in the expected direction.

Correlations between SGRQ scores and 6MWD, FVC % pred and gas exchange analysis are presented in tables 3, 4 and 5. Cross-sectional correlations between the SGRQ total, activity and impact scores and the 6MWD were moderate, but the correlation between 6MWD and the symptoms score was weak. All correlations between changes in SGRQ scores and 6MWD were weak. All correlations between SGRQ scores and FVC % pred or gas exchange analysis were weak except for correlations between changes in SGRQ scores and FVC % pred at week 52, which were moderate.

TABLE 3 Correlations (Spearman coefficients) between St George's Respiratory Questionnaire (SGRQ) scores and 6-min walk distance (6MWD)

SGRQ score	6MWD		
	Baseline	Change from baseline to week 24	Change from baseline to week 52
Total	−0.34***	−0.21***	−0.17**
Activity domain	−0.35***	−0.18***	−0.17**
Impact domain	−0.30***	−0.15**	−0.16**
Symptoms domain	−0.25***	−0.20***	−0.09

**: $p < 0.01$; ***: $p < 0.001$.

TABLE 4 Correlations (Spearman coefficients) between St George's Respiratory Questionnaire (SGRQ) scores and forced vital capacity (FVC) % pred and gas exchange analysis at baseline

SGRQ score	FVC % pred	D_{LCO} % pred	S_{pO_2}	P_{aO_2}	P_{aCO_2}
Total	−0.15**	−0.23***	−0.12*	−0.17***	−0.05
Activity domain	−0.11*	−0.26***	−0.09	−0.19***	−0.13*
Impact domain	−0.14**	−0.19**	−0.13*	−0.17***	−0.02
Symptoms domain	−0.14**	−0.15**	−0.06	−0.04	0.03

D_{LCO}: diffusing capacity of the lung for carbon monoxide; S_{pO_2}: arterial oxygen saturation measured by pulse oximetry; P_{aO_2}: arterial oxygen tension; P_{aCO_2}: arterial carbon dioxide tension. *: $p < 0.05$; **: $p < 0.01$; ***: $p < 0.001$.

TABLE 5 Correlations (Spearman coefficients) between changes in St George's Respiratory Questionnaire (SGRQ) scores and changes in forced vital capacity (FVC) % pred and gas exchange analysis at weeks 24 and 52

SGRQ score	D_{LCO} % pred		FVC % pred	
	Change from baseline to week 24	Change from baseline to week 52	Change from baseline to week 24	Change from baseline to week 52
Total	−0.20***	−0.19**	−0.29***	−0.38***
Activity domain	−0.15**	−0.16**	−0.21***	−0.35***
Impact domain	0.19***	−0.17**	−0.24***	−0.33***
Symptoms domain	−0.08	−0.12	−0.26***	−0.31***

D_{LCO}: diffusing capacity of the lung for carbon monoxide. **: $p < 0.01$; ***: $p < 0.001$.

Known-groups validity

Although there was overlap in the distribution of SGRQ scores between the subgroups defined by FVC % pred, mean SGRQ total score was higher in patients with more impaired lung function at baseline (FVC % pred ⩽70%) than in those with less impaired lung function at baseline (FVC % pred >70%) (figure 1). Similarly, although the distribution of SGRQ scores overlapped between the groups, mean SGRQ total score was higher in patients with D_{LCO} % pred ⩽55% than >55% at baseline (figure 2) and in patients with CPI >45 than ⩽45 at baseline (online supplementary figure S1).

Responsiveness

In general, there was a trend for greater changes in SGRQ total scores with greater absolute changes in FVC % pred, suggesting that SGRQ scores were sensitive to detecting change in patients whose health status declined or improved (figure 3 and online supplementary table S2). However, changes in SGRQ total score were only statistically significant *versus* the stable group for the subgroups showing deterioration in FVC % pred. The mean changes in SGRQ total score in patients with minimal or moderate deterioration in FVC % pred suggested that increases of 3–4 points over 52 weeks represented a meaningful deterioration (online supplementary table S2). No conclusion could be drawn about a meaningful improvement in SGRQ total score. There were significant differences between changes in SGRQ total score between patients reporting improvement or deterioration *versus* no change according to the MRC dyspnoea score (online supplementary table S3). The mean changes in SGRQ total score in patients with improvement or deterioration according to the MRC dyspnoea score suggested that changes of 8–9 points over 52 weeks represented meaningful improvement or deterioration (online supplementary table S3).

Discussion

In this study, we examined certain psychometric properties of the SGRQ in a cohort of patients with IPF who were participating in a clinical trial of a then investigational therapy. We found the SGRQ total score
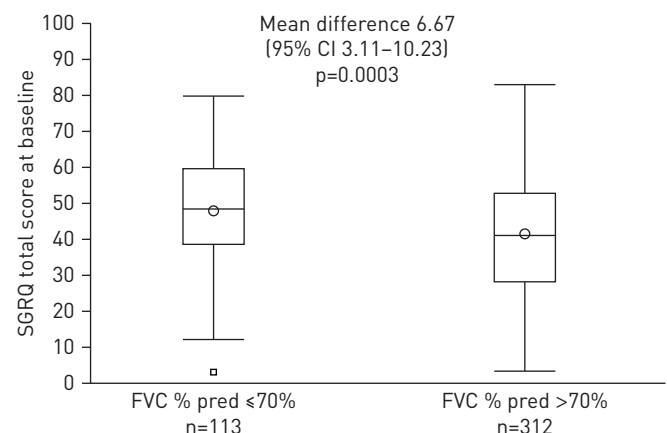

FIGURE 1 St George's Respiratory Questionnaire (SGRQ) total score by forced vital capacity (FVC) % pred category at baseline. The circles denote the mean values, the midline of the boxes indicate the median values, and boundaries denote 25th and 75th percentiles; whiskers are the minimal and maximum values in the lower fence (1.5 interquartile range above 75th percentile) and upper fence (1.5 interquartile range below 25th percentile). Outliers are depicted as boxes outside the whiskers.
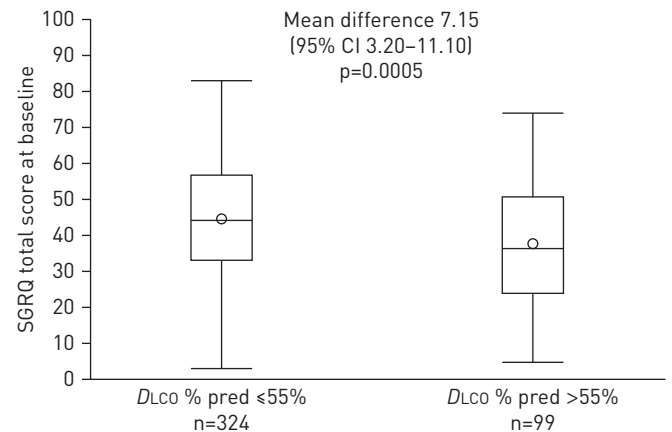
FIGURE 2 St George's Respiratory Questionnaire (SGRQ) total score by diffusing capacity of the lung for carbon monoxide (D_{LCO}) % pred category at baseline. The circles denote the mean values, the midline of the boxes indicate the median values, and boundaries denote 25th and 75th percentiles; whiskers are the minimum and maximum values in the lower fence (1.5 interquartile range above 75th percentile) and upper fence (1.5 interquartile range below 25th percentile).

to possess high internal consistency, suggesting relatively precise scores in this population. The test–retest reliability in patients with stable disease was acceptable, particularly given the 52-week trial duration. Tests of construct validity demonstrated weak to moderate correlations with four other measures of disease severity: a patient-reported outcome designed to provide a global perception of dyspnoea, a measure of exercise tolerance, FVC % pred and gas exchange analysis.

The finding that SGRQ total, activity and impact scores had high internal consistency, and that the symptoms scores had acceptable internal consistency, is in agreement with data from a randomised, placebo-controlled, phase III trial of another drug (bosentan) investigated as a potential therapy for IPF [17, 18]. As in the current study, these analyses reported high internal consistency for SGRQ total, activity and impact scores, but a lower internal consistency for the symptoms score. This is likely due to certain items within the domain lacking
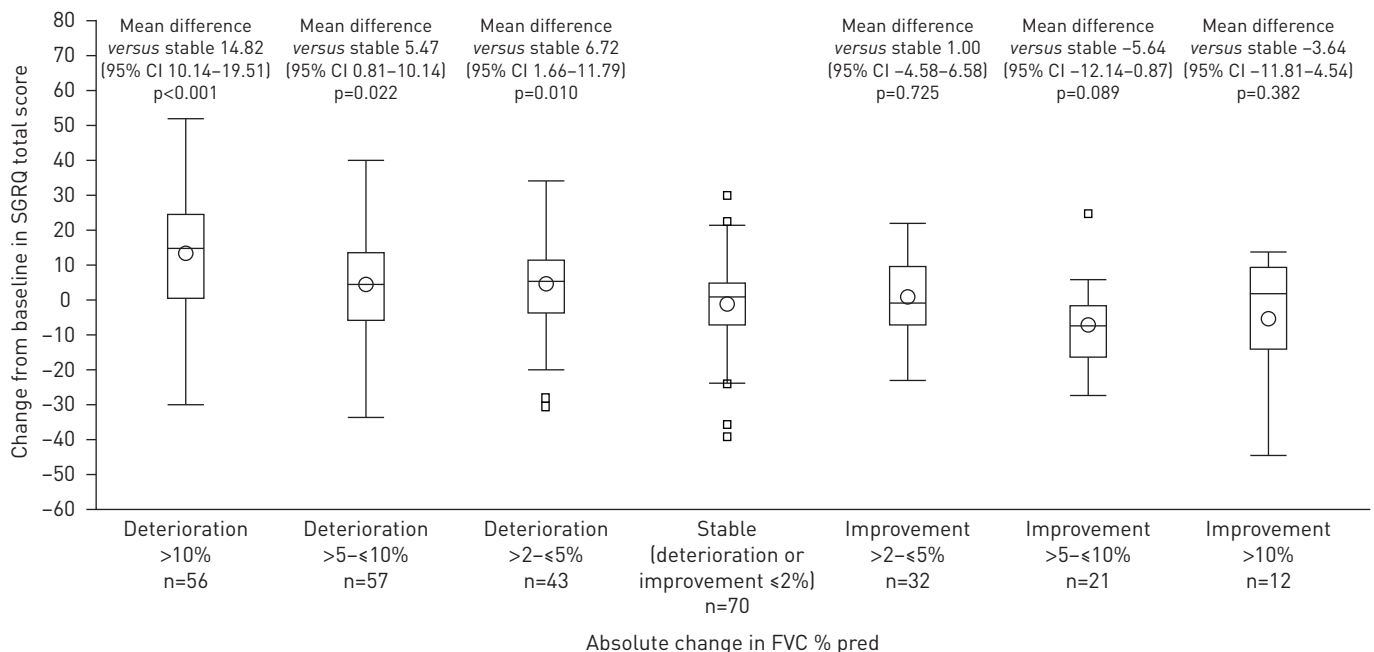

FIGURE 3 Change from baseline in St George's Respiratory Questionnaire (SGRQ) total score at week 52 by absolute change in forced vital capacity (FVC) % pred at week 52 (deterioration of >10%, deterioration of >5–⩽10%, deterioration of >2–⩽5%, deterioration or improvement of ⩽2% (stable), improvement of >2–⩽5%, improvement of >5–⩽10% and improvement of >10%). The circles denote the mean values, the midline of the boxes indicate the median values, and boundaries denote 25th and 75th percentiles; whiskers are the minimum and maximum values in the lower fence (1.5 interquartile range above 75th percentile) and upper fence (1.5 interquartile range below 25th percentile). Outliers are depicted as boxes outside the whiskers.

clinical relevance for patients with IPF. The major symptoms experienced by patients with IPF are cough and dyspnoea [5]; however, the symptoms domain examines a range of respiratory symptoms, many of which may apply to only a few patients with IPF. These off-target items may result in greater measurement error, reflected in the lower internal consistency of the symptoms domain, and weaken the association between its score and other clinical measures of disease severity in patients with IPF [8]. This could also introduce a conservative bias in clinical trials, making it more difficult to observe treatment effects. The limitations of the SGRQ in patients with IPF may partly be addressed by the use of IPF-specific questionnaires such as the SGRQ-I [18], but further data from prospective studies are needed to evaluate the specificity and responsiveness of such tools.

No data on the test–retest reliability (reproducibility) of the SGRQ in patients with stable IPF has previously been reported. In this study, SGRQ total scores were stable in patients whose FVC % pred did not change by >2% over 52 weeks. Although test–retest reliability is traditionally evaluated over a much shorter time period (e.g. 2 weeks), this analysis examined reproducibility over 52 weeks, which is more consistent with the typical duration of clinical trials in patients with IPF.

The validity of SGRQ total and domain scores was supported through moderate, statistically significant, cross-sectional correlations with the MRC dyspnoea score and 6MWD, which were in the hypothesised direction. Moderate correlations suggest that these tests are measuring related but different aspects of the symptomatology and impact of IPF. This is preferable to "perfect" correlations, which would suggest that the SGRQ is redundant as an assessment of health status, as other measures would suffice [4]. Several studies conducted in patients with IPF have reported moderate to strong correlations between SGRQ total and domain scores and pulmonary-specific patient-reported outcomes, including the Baseline Dyspnoea Index [19–21], the Cough Quality of Life Questionnaire [22], the King's Brief Interstitial Lung Disease Questionnaire [23] and the University of California San Diego Shortness of Breath Questionnaire [24]. Previous studies in patients with IPF have found that correlations with 6MWD tend to be moderate to strong for SGRQ total, activity and impact scores, and weak to moderate for the symptoms score [10, 18, 21, 25–27]. In this study, weak correlations were observed between SGRQ scores and pulmonary function tests. This is consistent with the results of other studies [18, 19, 21, 25, 26, 28] and, although causation cannot be inferred from correlation analyses, suggests that it is not only lung function that affects how a patient with IPF feels day to day.

There was also some evidence that the SGRQ was sensitive to detecting change in patients whose health status declined or improved according to change in FVC % pred and according to change in self-reported dyspnoea. These data support the results of a previous study in which investigators observed that changes in SGRQ score were sensitive to changes in disease status in patients with IPF over 6 months [17].

Limitations of this study include that these analyses were *post hoc*. Further assessment of the SGRQ in patients with IPF is needed. Generating prospective data on the psychometric properties of the SGRQ in larger clinical trials will allow for a more accurate assessment of its utility in this patient population. Additional data on the responsiveness of the SGRQ in patients with IPF are needed to inform interpretation of the scores in patients whose pulmonary function deteriorates or improves over time. Based on the eligibility criteria for the TOMORROW trial, few patients with severe lung function impairment were included in our analyses and more data are needed in such patients.

In conclusion, in a sample of 428 patients with IPF participating in a clinical trial, the SGRQ was found to perform well on tests of internal consistency, construct validity and reliability, and to be sensitive to detecting change. These results suggest that although it was not developed for use in patients with IPF, the SGRQ is an appropriate patient-reported outcome to measure HRQoL in clinical trials in patients with IPF.

References
1   Raghu G, Collard HR, Egan JJ, *et al.* An official ATS/ERS/JRS/ALAT statement: idiopathic pulmonary fibrosis: evidence-based guidelines for diagnosis and management. *Am J Respir Crit Care Med* 2011; 183: 788–824.
2   Nalysnyk L, Cid-Ruzafa J, Rotella P, *et al.* Incidence and prevalence of idiopathic pulmonary fibrosis: review of the literature. *Eur Respir Rev* 2012; 21: 355–361.
3   Ley B, Collard HR, King TE Jr. Clinical course and prediction of survival in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2011; 183: 431–440.
4   Swigris JJ, Kuschner WG, Jacobs SS, *et al.* Health-related quality of life in patients with idiopathic pulmonary fibrosis: a systematic review. *Thorax* 2005; 60: 588–594.

5   Swigris JJ, Stewart AL, Gould MK, *et al.* Patients' perspectives on how idiopathic pulmonary fibrosis affects the quality of their lives. *Health Qual Life Outcomes* 2005; 3: 61.

6   Belkin A, Swigris JJ. Health-related quality of life in idiopathic pulmonary fibrosis: where are we now? *Curr Opin Pulm Med* 2013; 19: 474–479.

7   Jones PW, Quirk FH, Baveystock CM. The St George's Respiratory Questionnaire. *Respir Med* 1991; 85: Suppl. B, 25–31.

8   Swigris JJ, Esser D, Conoscenti CS, *et al.* The psychometric properties of the St George's Respiratory Questionnaire (SGRQ) in patients with idiopathic pulmonary fibrosis: a literature review. *Health Qual Life Outcomes* 2014; 12: 124.

9   Richeldi L, Costabel U, Selman M, *et al.* Efficacy of a tyrosine kinase inhibitor in idiopathic pulmonary fibrosis. *N Engl J Med* 2011; 365: 1079–1087.

10   du Bois RM, 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: 1231–1237.

11   ATS Committee on Proficiency Standards for Clinical Pulmonary Function Laboratories. ATS statement: guidelines for the six-minute walk test. *Am J Respir Crit Care Med* 2002; 166: 111–117.

12   Fletcher CM, Elmes PC, Fairbairn AS, *et al.* The significance of respiratory symptoms and the diagnosis of chronic bronchitis in a working population. *Br Med J* 1959; 2: 257–266.

13   Nunnally JC, Bernstein IH. Psychometric Theory. 3rd Edn. New York, McGraw-Hill, 1994.

14   du Bois RM, 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: 1382–1389.

15   Hays R, Revicki DA. Reliability and validity (including responsiveness). *In*: Fayers P, Hays R, eds. Assessing Quality of Life in Clinical Trials. 2nd Edn. New York, Oxford University Press, 2005; pp. 25–39.

16   Hinkle D, Wiersma W, Jurs S. Applied Statistics for the Behavioral Sciences. 4th Edn. Boston, Houghton-Mifflin, 1988.

17   Swigris JJ, Brown KK, Behr J, *et al.* The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respir Med* 2010; 104: 296–304.

18   Yorke J, Jones PW, Swigris JJ. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. *Thorax* 2010; 65: 921–926.

19   Nishiyama O, Taniguchi H, Kondoh Y, *et al.* Health-related quality of life in patients with idiopathic pulmonary fibrosis. What is the main contributing factor? *Respir Med* 2005; 99: 408–414.

20   Yorke J, Swigris J, Russell AM, *et al.* Dyspnea-12 is a valid and reliable measure of breathlessness in patients with interstitial lung disease. *Chest* 2011; 139: 159–164.

21   Zimmermann CS, Carvalho CR, Silveira KR, *et al.* Comparison of two questionnaires which measure the health-related quality of life of idiopathic pulmonary fibrosis patients. *Braz J Med Biol Res* 2007; 40: 179–187.

22   Lechtzin N, Hilliard ME, Horton MR. Validation of the Cough Quality-of-Life Questionnaire in patients with idiopathic pulmonary fibrosis. *Chest* 2013; 143: 1745–1749.

23   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: 804–810.

24   Swigris JJ, Han M, Vij R, *et al.* The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respir Med* 2012; 106: 1447–1455.

25   Chang JA, Curtis JR, Patrick DL, *et al.* Assessment of health-related quality of life in patients with interstitial lung disease. *Chest* 1999; 116: 1175–1182.

26   Peng S, Li Z, Kang J, *et al.* Cross-sectional and longitudinal construct validity of the Saint George's Respiratory Questionnaire in patients with IPF. *Respirology* 2008; 13: 871–879.

27   Verma G, Marras T, Chowdhury N, *et al.* Health-related quality of life and 6 min walk distance in patients with idiopathic pulmonary fibrosis. *Can Respir J* 2011; 18: 283–287.

28   Tzanakis N, Samiou M, Lambiri I, *et al.* Evaluation of health-related quality-of-life and dyspnea scales in patients with idiopathic pulmonary fibrosis. Correlation with pulmonary function tests. *Eur J Intern Med* 2005; 16: 105–112.