Validation of a derived version of the IPF-specific Saint George's Respiratory Questionnaire

Thomas Skovhus Prior[1]* , Nils Hoyer[2], Saher Burhan Shaker[2], Jesper Rømhild Davidsen[3], Ole Hilberg[4], Haridarshan Patel[5] and Elisabeth Bendstrup[1]

Abstract

Background: Health-related quality of life (HRQL) is impaired in patients with idiopathic pulmonary fibrosis (IPF). HRQL is often measured using the St. George's Respiratory Questionnaire (SGRQ) despite the development of an IPF-specific version (SGRQ-I). Using data from a real-world cohort of patients with IPF, we aimed to transform SGRQ into a derived version of SGRQ-I, SGRQ-I$_{der}$, to examine the cross-sectional and longitudinal validity of SGRQ-I$_{der}$ and to compare SGRQ-I$_{der}$ to SGRQ-I.

Methods: Based on results from SGRQ, SGRQ-I$_{der}$ was derived applying the algorithm used to develop SGRQ-I. Of the 50 items in SGRQ, 34 items were retained in SGRQ-I$_{der}$. Response options for seven items were collapsed and minor adjustments were made to the weights of two items after correspondence with the developers of SGRQ-I. Cross-sectional validation, responsiveness and minimal clinically important difference (MCID) were assessed by comparison to other HRQL instruments, pulmonary function tests and 6-min walk test performed at baseline, 6 and 12 months. Furthermore, the association between SGRQ-I$_{der}$ scores and mortality was examined.

Results: A total of 150 IPF patients participated and 124 completed follow-up at 12 months. SGRQ-I$_{der}$ performed comparably to SGRQ-I with a high concurrent validity, good test–retest reliability and high known-groups validity. SGRQ-I$_{der}$ was responsive to change in HRQL and physiological anchors. MCID of SGRQ-I$_{der}$ for improvement and deterioration was 3.5 and 5.7, respectively. SGRQ-I$_{der}$ scores were associated with mortality in both univariate (HR 1.82, 95% CI 1.42–2.34 per 20-point increase) and multivariate analyses (HR 1.57, 95% CI 1.20–2.05 per 20-point increase).

Conclusions: The SGRQ-I$_{der}$ is a valid, reliable and responsive HRQL instrument in patients with IPF and has psychometric properties comparable to SGRQ-I. Thus, SGRQ results can reliably be transformed into the SGRQ-I$_{der}$. The MCID estimates were calculated for improvement and deterioration separately. Increasing SGRQ-I$_{der}$ score was associated with increased mortality.

Keywords: Idiopathic pulmonary fibrosis, Quality of life, Validation, St. George's Respiratory Questionnaire, IPF-specific version of St. George's Respiratory Questionnaire

Introduction

Idiopathic pulmonary fibrosis (IPF) is a chronic fibrotic lung disease with a large respiratory symptom burden and poor prognosis [1]. Dyspnea, cough, fatigue, social isolation, loss of emotional well-being and numerous comorbidities lead to impaired health-related quality

*Correspondence: thbjer@rm.dk
[1] Center for Rare Lung Diseases, Department of Respiratory Diseases and Allergy, Aarhus University Hospital, Aarhus, Denmark
Full list of author information is available at the end of the article

Prior *et al. Respir Res*　(2021) 22:259

Page 2 of 10

of life (HRQL) in patients with IPF [2–5]. As HRQL is regarded as an increasingly important outcome, both in clinical trials and daily clinical practice, valid and reliable HRQL measures are required [6].

Various instruments are used to measure HRQL in patients with IPF. Due to the lack of IPF-specific HRQL instruments, the St. George's Respiratory Questionnaire (SGRQ) has often been used in IPF studies even though it was developed for patients with obstructive lung diseases [7, 8]. SGRQ has adequate psychometric properties in IPF, but patients with obstructive lung diseases have a different symptom profile, and some items in SGRQ are less relevant to patients with IPF. Hence, especially the symptoms domain of SGRQ has weaker psychometric properties in IPF [8].

To meet these drawbacks, an IPF-specific version of SGRQ (SGRQ-I) was developed and validated [5, 9, 10]. Only the items most relevant to patients with IPF were retained, resulting in a 34-item instrument compared to the 50 items in SGRQ. However, SGRQ-I has not been widely adopted and SGRQ is still used broadly in clinical trials. The assessment of HRQL could probably be more specific to patients with IPF, if the results from clinical trials using SGRQ, instead of SGRQ-I, could be transformed into an equally valid and reliable derived version of SGRQ-I (SGRQ-I_{der}). The SGRQ-I_{der} should be validated both cross-sectionally and longitudinally to ensure solid psychometric properties.

Using data from a real-world cohort of patients with IPF, the aim of this study was to transform SGRQ into a derived version of SGRQ-I (SGRQ-I_{der}), to examine the cross-sectional and longitudinal validity of SGRQ-I_{der} and to compare SGRQ-I_{der} to SGRQ-I. Furthermore, we aimed to determine the minimal clinically important difference (MCID) for SGRQ-I_{der} and to examine the ability of SGRQ-I_{der} to predict mortality in patients with IPF.

Materials and methods
Study subjects
The current study was based on the cohort previously used for cross-sectional and longitudinal validation of SGRQ-I and the King's Brief Interstitial Lung Disease questionnaire (K-BILD) as well as comorbidities in IPF [4, 5, 10, 11]. Adult patients with a guideline-based diagnosis of IPF and attending one of the tertiary interstitial lung disease (ILD) centers in Denmark at the university hospitals in Aarhus, Copenhagen and Odense were eligible for inclusion [18, 19]. Inability to complete the instruments due to linguistic or cognitive barriers excluded patients from participation.

Study measures
Based on results from SGRQ, SGRQ-I_{der} was derived applying the algorithm used to develop SGRQ-I [9]. The derivation algorithm was determined by two separate authors (TSP and HP) and compared; in case of disagreement, consensus was obtained in consultation with EB. Of the 50 items in SGRQ, the 34 items most relevant to patients with IPF were retained in SGRQ-I_{der}. Response options for seven items were collapsed (all items in the symptoms domain and the first item in the activities domain, Additional file 1: Table S1). Minor adjustments were made to the weights of two items compared to the SGRQ-I algorithm after correspondence with the developers of SGRQ-I, Jeff Swigris and Janelle Yorke (reverse scoring in the last item in the symptoms domain and a minor correction to the weight of the first item in the activities domain, Additional file 1: Table S2).

At baseline, SGRQ was completed along with SGRQ-I, K-BILD, Short Form-36 (SF-36) and University of California San Diego Shortness of Breath questionnaire (SOBQ). Global Rating of Change Scales (GRCS) and SGRQ were completed 14 days later. Patients completed GRCS, SGRQ, SGRQ-I, K-BILD, and SOBQ after 6 and 12 months.

Pulmonary function tests (PFT) were performed to assess forced vital capacity (FVC) and diffusing capacity of the lung for carbon monoxide (DLCO) along with a 6-min walk test (6MWT) at baseline, 6 and 12 months. Based on this information, the gender, age and physiology (GAP) index was calculated.

Comorbidities were registered at baseline by review of patients' medical history, indications for medications, blood samples, echocardiography and chest high-resolution computed tomography (HRCT) scans, and the Charlson comorbidity index was calculated.

SGRQ-I contains 34 items measuring HRQL and has been cross-sectionally and longitudinally validated for use in IPF [5, 9, 10]. SGRQ-I scores were divided into deteriorated (ΔSGRQ-I $\geq$ 4.9), unchanged (ΔSGRQ-I − 3.9 to 4.9) and improved (ΔSGRQ-I $\leq$ − 3.9) in accordance with the MCIDs for SGRQ-I [10].

K-BILD is a HRQL instrument developed for patients with ILD and validated in IPF [4, 10, 12]. Based on the MCIDs for K-BILD, patients were classified as deteriorated (ΔK-BILD $\leq$ − 2.7), unchanged (ΔK-BILD − 2.7 to 4.7) and improved (ΔK-BILD $\geq$ 4.7) [10].

SOBQ measures dyspnea related to daily activities [13]. SOBQ has both cross-sectional and longitudinal validity in IPF [14]. In accordance with the MCID of SOBQ in IPF, changes in SOBQ score were classified as deteriorated (ΔSOBQ $\geq$ 8), unchanged (ΔSOBQ − 8 to 8) or improved (ΔSOBQ $\leq$ − 8) [14].

Prior *et al. Respir Res*   (2021) 22:259

Page 3 of 10

SF-36 is a generic instrument assessing a range of quality of life domains [15].

GRCS ranging from −5 "Very much worse" over 0 "Unchanged" to 5 "Very much better" were used to estimate changes from baseline in HRQL [16]. Four specific GRCS were composed: one to reflect the overall HRQL and three for the domains of SGRQ-I$_{der}$. Results from each GRCS was categorized as deteriorated (−5 to −2), unchanged (−1 to 1) or improved (2 to 5).

FVC and *DLCO* are commonly used in IPF as indicators of disease severity, and functional capacity can be evaluated by the distance walked during the 6-min walk test (*6MWD*). Mortality in patients with IPF is associated with both PFTs and 6MWD [17, 18]. Based on the MCID for FVC in ILD, patients with an absolute change in FVC % predicted below 6% were regarded as unchanged, whereas an absolute change larger than or equal to ±6% was regarded as improved or deteriorated, respectively [19, 20].

As no MCID has been reported for DLCO in IPF, the intraindividual variability in DLCO measurements was used [21]. Absolute changes in DLCO % predicted were divided into deteriorated (ΔDLCO ≤ −10%), unchanged (ΔDLCO −10% to 10%) and improved (ΔDLCO ≥ 10). Changes in 6MWD were divided into deteriorated (Δ6MWD ≤ −28 m), unchanged (Δ6MWD −28 m to 28 m) and improved (Δ6MWD ≥ 28 m), in concordance with the MCID for 6MWD in IPF [22].

The GAP index was developed to predict mortality in patients with IPF [23]. The index is calculated based on gender, age, FVC and DLCO. The resulting three groups have 1-year mortalities ranging from 6 to 39%.

Statistical analyses
Instruments with more than 15% missing answers or missing either total or domain scores were excluded from the analyses.

Cross-sectional validation
Baseline results were used to perform the cross-sectional validation of SGRQ-I$_{der}$. Floor and ceiling effects were defined by > 15% of patients scoring the highest or lowest possible scores, respectively. Difference between SGRQ-I and SGRQ-I$_{der}$ total and domain scores were assessed by the paired two sample t-test.

The internal consistency measures the interrelatedness of items in an instrument. To examine the internal consistency of SGRQ-I$_{der}$, Cronbach's α was calculated for the total and each domain score. Results > 0.7 were compatible with a reliable internal consistency [24].

Concurrent validity was evaluated by comparing SGRQ-I$_{der}$ to SGRQ-I, K-BILD, SOBQ, SF-36, PFTs and 6MWD. Intraclass correlation coefficients (ICCs) (2,1)

and Bland–Altman plots were used to compare SGRQ-I$_{der}$ to SGRQ-I, and SGRQ-I$_{der}$ was compared to the other measures using Pearson's correlation coefficients.

By comparing the scores at baseline and after two weeks in stable patients (scoring −1 to 1 in GRCS 2 weeks after baseline), test–retest reliability of SGRQ-I$_{der}$ was evaluated by ICCs (2,1) and a Bland–Altman plot. ICCs > 0.7 were regarded as a measure of reliability [24].

Known-groups validity was assessed by evaluating the ability of SGRQ-I$_{der}$ to differentiate between patients with different stages of disease severity. Patients were divided into "known groups" of disease severity according to 6MWD and PFTs (upper and lower quartiles), use of long-term oxygen therapy (LTOT) and GAP index [25]. The independent two-sample t-test was used for normally distributed continuous data and the Wilcoxon–Mann–Whitney test for not normally distributed continuous data. GAP groups were compared by linear regression analyses.

Data were analyzed using STATA, version 14.2 (Stata-Corp, College Station, Texas).

Longitudinal validation
Change in SGRQ-I$_{der}$ total score was analyzed using a mixed effects model with cluster effect for center (using the "Clustered Sandwich Estimator") and random intercept.

Responsiveness was assessed using Pearson's correlation coefficients to examine the association between changes in SGRQ-I$_{der}$ and changes in anchors (GRCS, SGRQ-I, K-BILD, SOBQ, FVC % predicted, DLCO % predicted, 6MWD) from baseline to 12 months. Negative correlations between SGRQ-I$_{der}$ and GRCS, K-BILD, FVC, DLCO and 6MWD were expected because of inverse scoring algorithms. Subgroup analyses to evaluate the effect of receiving antifibrotic treatment at baseline were performed.

The association between SGRQ-I$_{der}$ baseline score stratified into 20-point intervals and mortality for up to 48 months of follow-up was assessed using Cox regression analyses. Subsequently, the model was adjusted for age, FVC % predicted and the Charlson comorbidity index.

The MCID SGRQ-I$_{der}$ was estimated by receiver operating characteristic (ROC) curves. To estimate MCID, both anchor-based and distribution-based methods are recommended [26, 27], and a combination of these methods are included in ROC curves [28]. A correlation coefficient > 0.3 between SGRQ-I$_{der}$ and anchors to be included in the MCID analyses was prespecified, as anchors and the instrument under investigation must be related [26]. Based on thresholds of the anchors (described above),

Prior *et al. Respir Res*   (2021) 22:259

Page 4 of 10

patients were categorized as deteriorated, unchanged or improved. Separate ROC curves were used to estimate the MCID for deterioration (unchanged vs. deteriorated patients) and improvement (unchanged vs. improved patients) [28]. The optimal cut-off point of the ROC curve (with equal sensitivity and specificity) was regarded as the MCID estimate for each anchor. To assess the influence of antifibrotic therapy at baseline on MCID estimates, subgroup analyses were performed by similar ROC curve analyses.

Results

The study population consisted of 150 patients with IPF included from August 2016 to March 2018 (Table 1). The cohort was dominated by male patients receiving antifibrotic therapy at baseline with a history of smoking. DLCO was moderately reduced whereas FVC was relatively well preserved. Most patients completed the 6-month visit (n = 135, 90%) and the 12-month visit (n = 124, 83%). Patients were withdrawn from the study due to death (n = 16), inability to complete the instruments (n = 1), inability to attend the outpatient clinic (n = 3) or patient's wish to withdraw (n = 6).

Cross-sectional validation

No floor or ceiling effects in SGRQ-I_{der} total or domain scores were present. Only minor differences in item, total and domain scores between SGRQ-I_{der} and SGRQ-I were observed (Table 2 and Additional file 1: Table S3). A good internal consistency was indicated by high Cronbach's α results in both domain and total scores (Table 2).

A high concurrent validity was demonstrated by high ICCs and Bland–Altman plots comparing SGRQ-I_{der} to SGRQ-I (Table 2 and Fig. 1). These results were supported by moderate to strong correlations with K-BILD, SOBQ and SF-36 and weaker correlations with PFTs and 6MWD (Table 3 and Additional file 1: Table S4).

Most patients remained in a stable health status after 14 days as indicated by GRCS. Based on results from the stable patients, high ICCs and a Bland–Altman plot indicated a good test–retest reliability of SGRQ-I_{der}, which was comparable to SGRQ-I (Table 4, Fig. 2 and Additional file 1: Table S5).

The known-groups validity was high, as SGRQ-I_{der} demonstrated significantly better HRQL in patients with the highest quartiles of PFTs and 6MWD compared with the lowest quartiles. Better HRQL was also seen in patients without LTOT and decreasing HRQL was associated with increasing disease severity as indicated by the GAP index (Fig. 3).

Longitudinal validation

After 12 months, a non-significant HRQL decrease was observed (ΔSGRQ-I_{der} total: 2.7, 95% CI − 0.4 to 7.8). According to the HRQL anchors, 16–33% of patients improved and 20–44% deteriorated, whereas 2–20% improved and 16–35% deteriorated according to the physiological anchors.

Table 1 Baseline characteristics

Characteristics	Value
Total cohort, n	150
Male gender (%)	122 (81.3%)
Age, years (SD)	72.9 ± 6.2
Smoking status	
Never (%)	40 (26.6%)
Former (%)	101 (67.3%)
Current (%)	9 (6.0%)
FVC, % predicted (SD)	87.2 ± 23.1
DLCO, % predicted (SD)	48.4 ± 14.1
6MWD, m (SD)	450.3 ± 112.5
Concurrent long-term oxygen therapy (%)	19 (12.7%)
Antifibrotic treatment, n (%)	85 (56.7%)
Pirfenidone, n (%)	51 (34.0%)
Nintedanib, n (%)	34 (22.7%)
Charlson comorbidity index (IQR)	1 (0–2)

Values are presented as *n* (%), mean ± standard deviation (SD), or median with interquartile range (IQR) [4, 5, 10, 11]

FVC: Forced vital capacity; DLCO: Diffusing capacity of the lung for carbon monoxide; 6MWD: distance walked during the 6-min walk test

Table 2 Summary scores, internal consistency and concurrent validity of SGRQ-I_{der}

SGRQ-I_{der}	SGRQ-I_{der} Mean (SD)	SGRQ-I Mean (SD)	Difference (95% CI)	Cronbach's α	ICC
Total	43.3 (21.5)	43.4 (22.1)	0.03 (− 0.45 to 0.51)	0.92	0.99
Symptoms	49.8 (22.9)	49.6 (27.1)	− 0.10 (− 2.74 to 2.53)	0.74	0.80
Activities	61.5 (28.3)	61.5 (28.3)	0.03 (0.02 to 0.03)	0.86	1.00
Impacts	32.0 (22.3)	31.9 (22.3)	0.12 (− 0.07 to 0.31)	0.84	1.00

Data are presented as mean with standard deviations (SD) or 95% confidence interval (CI) of SGRQ-I_{der} compared to SGRQ-I and Cronbach's alpha and intraclass correlation coefficients (ICC) of SGRQ-I_{der} for the total and three domain scores in all patients

SGRQ-I_{der}: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire

Fig. 1 Bland–Altman plot of the agreement between SGRQ-I$_{der}$ and SGRQ-I. The solid line represents the mean difference and dashed lines represent 95% limits of agreement. SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from th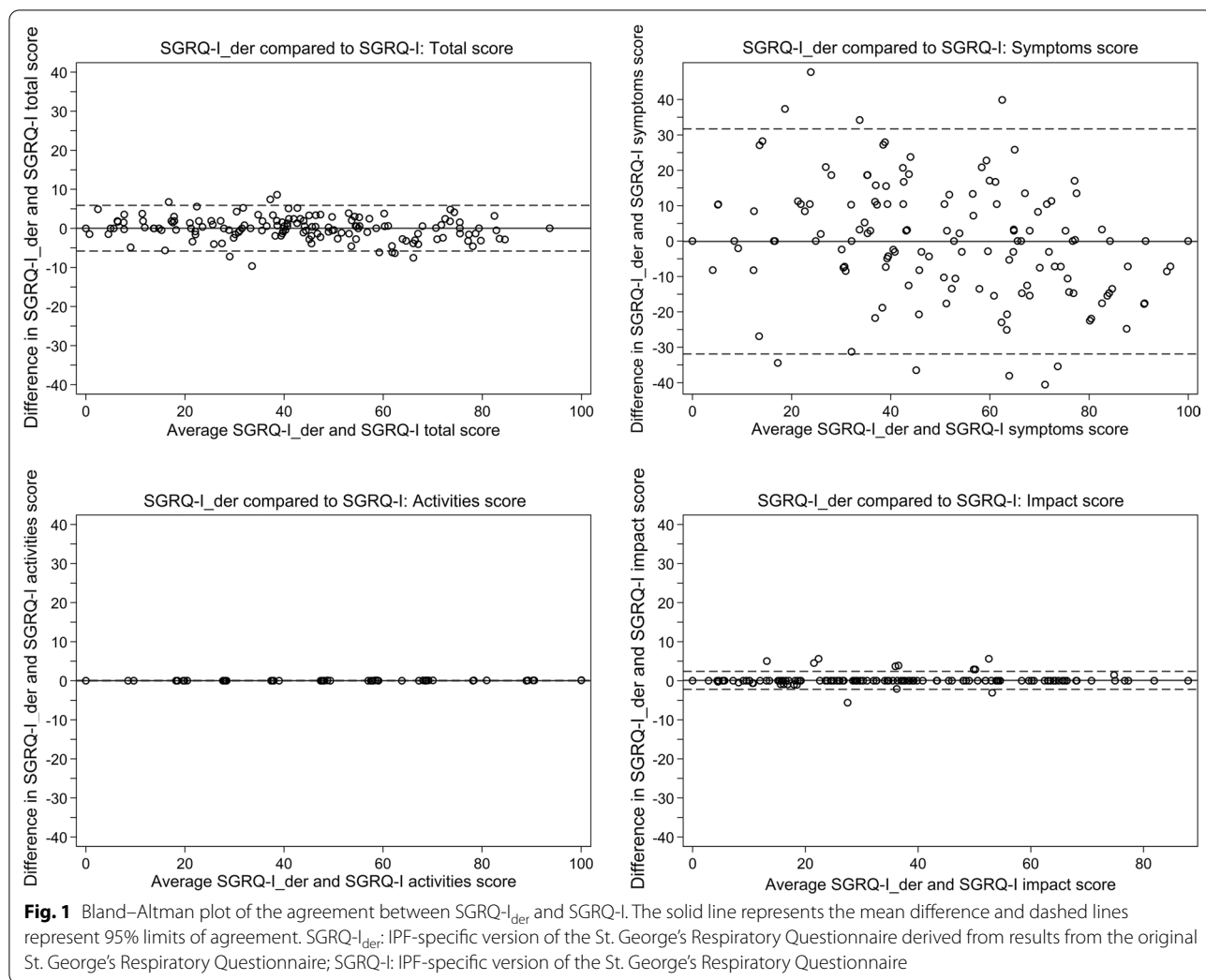e original St. George's Respiratory Questionnaire; SGRQ-I: IPF-specific version of the St. George's Respiratory Questionnaire

Table 3 Concurrent validity of SGRQ-I$_{der}$

	SGRQ-I$_{der}$ total	SGRQ-I$_{der}$ symptoms	SGRQ-I$_{der}$ activities	SGRQ-I$_{der}$ impacts
K-BILD total	− 0.76	− 0.57	− 0.71	− 0.70
K-BILD chest symptoms	− 0.68	− 0.58	− 0.54	− 0.67
K-BILD breathlessness and activities	− 0.78	− 0.58	− 0.76	− 0.70
K-BILD psychological	− 0.58	− 0.44	− 0.52	− 0.55
SOBQ total	0.80	0.52	0.73	0.77
SF-36 PCS	− 0.71	− 0.50	− 0.63	− 0.69
SF-36 MCS	− 0.46	− 0.40	− 0.34	− 0.46
FVC % predicted	− 0.29	− 0.23	− 0.20	− 0.30
DLCO % predicted	− 0.49	− 0.31	− 0.53	− 0.43
6MWD (m)	− 0.51	− 0.26	− 0.47	− 0.51

Data are presented as Pearson's correlation coefficients for all patients

SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire; SGRQ-I: IPF-specific version of the St. George's Respiratory Questionnaire; K-BILD: King's Brief Interstitial Lung Disease questionnaire; SOBQ: University of California, San Diego Shortness of Breath Questionnaire; SF-36: Short Form-36; PCS: Physical Component Score; MCS: Mental Component Score; FVC: Forced vital capacity; DLCO: Diffusing capacity of the lung for carbon monoxide; 6MWD: Distance walked during the 6-min walk test

Prior *et al. Respir Res*   (2021) 22:259

Page 6 of 10

Table 4 Test–retest reliability of SGRQ-I$_{der}$

SGRQ-I$_{der}$	n	ICC
Total	99 (73.9%)	0.91
Symptoms	105 (78.4%)	0.77
Activities	104 (77.6%)	0.80
Impacts	104 (77.6%)	0.76

Data represent number of stable patients (% of responders, $n = 134$) and intraclass correlation coefficients (ICCs) [5]

SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire

Fig. 2 Bland–Altman plot of the test–retest reliability of SGRQ-I$_{der}$ in all stable patients. The solid line represents the mean difference and dashed lines represent 95% limits of agreement. SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire
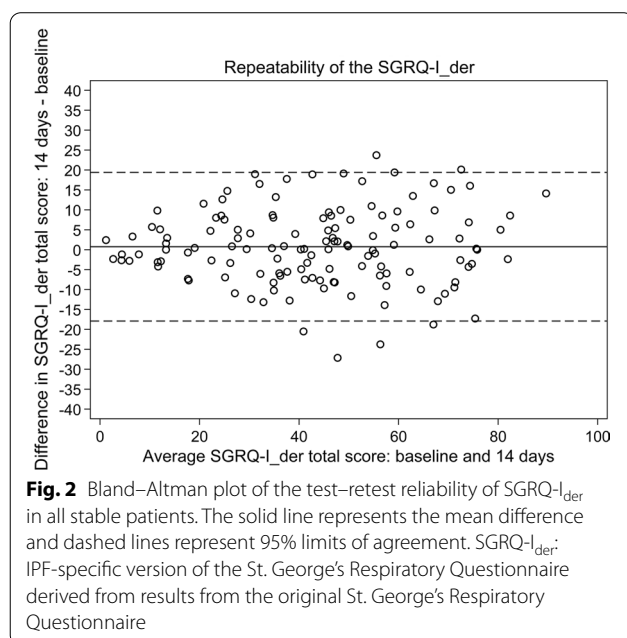
SGRQ-I$_{der}$ was responsive to changes in HRQL anchors and, to a lesser extent, also physiological anchors similar to SGRQ-I (Table 5 and Additional file 1: Table S6). Correlations for patients receiving antifibrotic therapy at baseline were comparable. All correlations were in the expected direction.

The MCID estimates for SGRQ-I$_{der}$ are presented in Table 6. Subgroup analyses for patients receiving antifibrotic treatment at baseline were comparable (SGRQ-I$_{der}$ total score: improvement 3.9, deterioration 6.5).

Increasing baseline SGRQ-I$_{der}$ total score in 20-point intervals was associated with increasing mortality in the univariate model (HR 1.82, 95% CI 1.42–2.34) and remained significant after adjustment for FVC % predicted, age and the Charlson comorbidity index (HR 1.57, 95% CI 1.20–2.05) (Fig. 4).

Discussion

This is the first study to directly transform SGRQ results from a real-world, multicenter cohort of patients with IPF into a derived version of SGRQ-I and to evaluate the validity and reliability of this HRQL instrument. Given the extensive use of SGRQ and limited utilization of SGRQ-I in IPF research, results may become suboptimal due to the content and psychometric properties of SGRQ being less specific for patients with IPF. Our study showed that SGRQ data can be transformed into SGRQ-I$_{der}$ with a validity, reliability and responsiveness comparable to SGRQ-I in patients with IPF. The mean total and domain scores of SGRQ-I$_{der}$ were almost identical to the mean scores of SGRQ-I, and as such performed comparably to SGRQ-I with only minor differences. MCID estimates for improvement and deterioration were estimated, facilitating the interpretation of repeated measurements. Furthermore, SGRQ-I$_{der}$ can be used to assess the prognosis of IPF.

The internal consistency of SGRQ-I$_{der}$ was high across both total score and the three domains, but the symptoms domain had the lowest performance. The same pattern was observed concerning concurrent validity. Generally, the symptoms domain had weaker correlations with both HRQL instruments and physiological parameters and in the Bland–Altman plots, the variation between SGRQ-I and SGRQ-I$_{der}$ exceeded the variation in the other domains. These differences could partly be explained by the changes made to the instrument during the development of SGRQ-I. In the symptoms domain, two items from SGRQ were removed, the response options in the remaining items were collapsed, and new weights were calculated for each response options. In the other domains, selected items were removed and response options and weights were mostly unchanged [9]. As the largest changes were made in the symptoms domain, the major differences between SGRQ-I and SGRQ-I$_{der}$ would be expected in this domain. Another explanation could be found in the inherent properties of the symptoms domain. The domain includes items concerning sputum, wheezing and attacks of chest trouble which are less relevant to patients with IPF. Therefore, the symptoms domain also has the weakest psychometric properties compared with the other domains of SGRQ and SGRQ-I when used in IPF populations [5, 8]. Despite these shortcomings, most items in the symptoms domain were preserved when developing the SGRQ-I, as the validity and reliability of the domain became weaker without the items and at the same time, the instrument would be better at assessing HRQL in IPF patients with comorbidities such as chronic obstructive pulmonary disease [9].
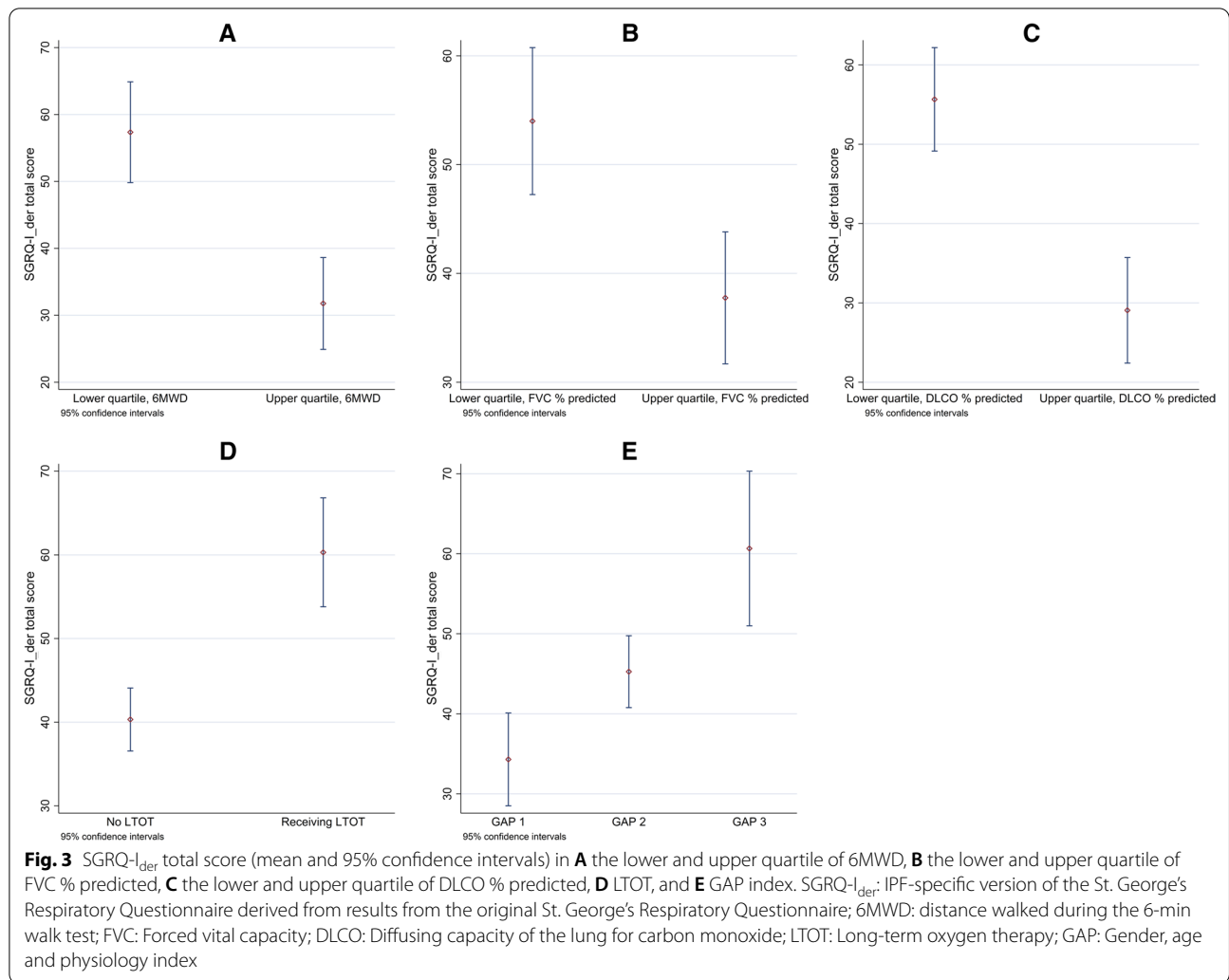
Prior *et al. Respir Res*   (2021) 22:259

Page 7 of 10

Fig. 3 SGRQ-I$_{der}$ total score (mean and 95% confidence intervals) in **A** the lower and upper quartile of 6MWD, **B** the lower and upper quartile of FVC % predicted, **C** the lower and upper quartile of DLCO % predicted, **D** LTOT, and **E** GAP index. SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire; 6MWD: distance walked during the 6-min walk test; FVC: Forced vital capacity; DLCO: Diffusing capacity of the lung for carbon monoxide; LTOT: Long-term oxygen therapy; GAP: Gender, age and physiology index

Table 5 Correlations between changes in SGRQ-I$_{der}$ domains and changes in anchors from baseline to 12 months

	GRCS	ΔSGRQ-I	ΔK-BILD	ΔSOBQ	ΔFVC%	ΔDLCO%	Δ6MWD
ΔSGRQ-I$_{der}$ total	− 0.55	0.95	− 0.56	0.61	− 0.24	− 0.24	− 0.48
ΔSGRQ-I$_{der}$ symptoms	− 0.51	0.62	− 0.45	0.46	− 0.27	− 0.23	− 0.38
ΔSGRQ-I$_{der}$ activities	− 0.47	0.68	− 0.48	0.44	− 0.23	− 0.21	− 0.29
ΔSGRQ-I$_{der}$ impacts	− 0.46	0.87	− 0.44	0.53	− 0.13	− 0.17	− 0.45

Δ: Change from baseline to 12 months; SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire; GRCS: Global rating of change scales; SGRQ-I: IPF-specific version of the St. George's Respiratory Questionnaire; K-BILD: King's Brief Interstitial Lung Disease questionnaire; SOBQ: University of California San Diego Shortness of Breath questionnaire; FVC%: Forced vital capacity % predicted; DLCO%: Diffusing capacity of the lung for carbon monoxide % predicted; 6MWD: Distance walked during the 6-min walk test

MCID estimates for SGRQ-I$_{der}$ were comparable to MCIDs for SGRQ-I (improvement 3.5 vs. 3.9 and deterioration 5.7 vs. 4.9, respectively) [10]. The small differences are probably caused by the minor deviations that exist between the two versions of the instrument. MCID estimates for SGRQ in other IPF studies were

slightly higher ranging from 4 to 6.6 [25, 29]. The divergence could be caused by a single MCID for deterioration and improvement combined, by characteristics of the cohorts and the statistical methods used. The SGRQ studies were based on clinical trial populations, whereas the present study was based on a real-world

Prior *et al. Respir Res*   (2021) 22:259

Page 8 of 10

Table 6 Mean and range of MCID estimates for SGRQ-I$_{der}$ domains based on change in anchors from baseline to 12 months

Domains	Improvement		Deterioration	
	Mean	Range	Mean	Range
SGRQ-I$_{der}$ total	3.5	2.9–4.5	5.7	4.7–6.6
SGRQ-I$_{der}$ symptoms	7.9	7.5–8.2	3.9	1.7–5.3
SGRQ-I$_{der}$ activities	10.2	10.0–10.9	9.6	9.5–9.6
SGRQ-I$_{der}$ impacts	1.2	0.3–2.8	4.1	2.3–5.8

MCID: Minimal clinically important difference; SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire

Fig. 4 Survival estimates for groups of patients with different SGRQ-I$_{der}$ total score. SGRQ-I$_{der}$: IPF-specific version of the St. George's Respiratory Questionnaire derived from results from the original St. George's Respiratory Questionnaire
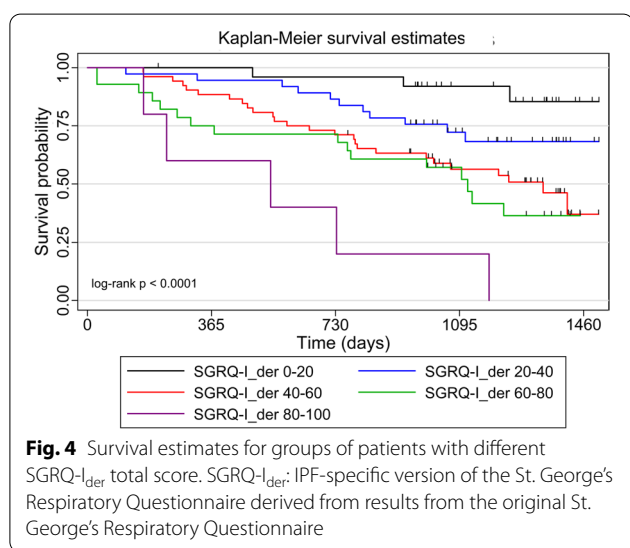
cohort of patients. Furthermore, more distribution-based approaches were used in these studies which tends to produce higher MCID estimates [30]. When using ROC curves, both anchor- and distribution-based methods are included in the model as recommended [28].

There was a clear association between decreasing baseline HRQL and increasing mortality, even after adjustment for covariates. The relationship between HRQL and mortality in IPF has been studied using SGRQ, but results are divergent. In two studies, baseline HRQL was shown to be significantly associated with mortality [31, 32], whereas two other studies did not find a significant association in multivariate analyses [33, 34]. Kreuter et al. found a significant association between HRQL at last available follow-up and mortality, but no association with change in HRQL from baseline [35]. Hence, it is possible that the IPF-specific versions of SGRQ are superior at predicting mortality due to the disease-specific nature of the instruments.

A strength of the present study is the inclusion of a large and broad, multicenter, real-world cohort of patients with IPF based on wide inclusion and few exclusion criteria. This increases the external validity of the results. Furthermore, the other HRQL instruments used for the analyses were validated for use in IPF, thus increasing the reliability of the results. A limitation is the possible recall bias related to GRCS used in the longitudinal analyses. Even though it can be difficult to recall ones health status 12 months ago, GRCS provide a simple evaluation of patients' HRQL and can be tailored to reflect both overall HRQL and domains of a HRQL instrument. In addition, GCRS have good validity, reliability and response to change over time [16]. Another potential limitation is healthy volunteer bias, as healthier patients may be more willing to participate in clinical trials [36]. This may limit the generalizability of the results, but the current cohort also included patients with advanced disease, thus limiting this type of bias.

In conclusion, a derived version of SGRQ-I transformed from SGRQ data, the SGRQ-I$_{der}$, is a valid, reliable and responsive HRQL instrument in patients with IPF and has psychometric properties comparable to SGRQ-I. Thus, SGRQ results can reliably be transformed into the SGRQ-I$_{der}$. The MCID estimate for improvement is 3.5 and 5.7 for deterioration, and increasing SGRQ-I$_{der}$ scores are associated with increased mortality.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12931-021-01853-2.

Additional file 1: Table S1. The SGRQ-I development algorithm. **Table S2.** Changes from the original scoring algorithm for SGRQ-I used in SGRQ-I$_{der}$. **Table S3.** Mean (SD) item scores at baseline. **Table S4.** Concurrent validity of SGRQ-I$_{der}$ and SGRQ-I. **Table S5.** Test-retest validity of SGRQ-I$_{der}$ and SGRQ-I. **Table S6.** Responsiveness of SGRQ-I$_{der}$ and SGRQ-I

Prior *et al. Respir Res*   (2021) 22:259

Page 9 of 10

Availability of data and materials

Declarations

Ethics approval and consent to participate

Written informed consent was obtained from all participants before inclusion in the study. The study was approved by the Central Denmark Region Committee on Health Research Ethics (Case no. 1-10-72-87-16) and registered at clinicaltrials.org (NCT02818712) before initiation.

Competing interests

For the current work, TSP has received an unrestricted grant from Boehringer Ingelheim and TSP and HP have received personal fees from Galapagos. Outside the current work, TSP has received personal fees from Boehringer Ingelheim, JRD has received personal fees and non-financial support from Boehringer Ingelheim and Roche, and EB has received personal fees from Galapagos and personal fees and grants from Boehringer Ingelheim and Roche. NH, OH, and SBS have nothing to declare.

Author details

[1]Center for Rare Lung Diseases, Department of Respiratory Diseases and Allergy, Aarhus University Hospital, Aarhus, Denmark. [2]Department of Respiratory Medicine, Herlev-Gentofte University Hospital, Copenhagen, Denmark. [3]South Danish Center for Interstitial Lung Diseases (SCILS), Department of Respiratory Medicine, Odense University Hospital, Odense, Denmark. [4]Department of Respiratory Medicine, Vejle Hospital, Vejle, Denmark. [5]Immensity Consulting, Inc., Chicago, IL, USA.

References

1. Raghu G, Chen S-Y, Yeh W-S, Maroni B, Li Q, Lee Y-C, et al. Idiopathic pulmonary fibrosis in US Medicare beneficiaries aged 65 years and older: incidence, prevalence, and survival, 2001–11. Lancet Respir Med. 2014;2600:1–7.
2. King CS, Nathan SD. Idiopathic pulmonary fibrosis: effects and optimal management of comorbidities. Lancet Respir Med. 2017;5:72–84.
3. Belkin A, Swigris JJ. Health-related quality of life in idiopathic pulmonary fibrosis: where are we now? Curr Opin Pulm Med. 2013;19:474–9.
4. Prior TS, Hilberg O, Shaker SB, Davidsen JR, Hoyer N, Birring SS, et al. Validation of the King's brief interstitial lung disease questionnaire in idiopathic pulmonary fibrosis. BMC Pulm Med. 2019;19:255. https://doi.org/10.1186/s12890-019-1018-0.
5. Prior TS, Hoyer N, Shaker SB, Davidsen JR, Yorke J, Hilberg O, et al. Validation of the IPF-specific version of St. George's Respiratory Questionnaire. Respir Res. 2019;20:199. https://doi.org/10.1186/s12931-019-1169-9.
6. Russell A-M, Sprangers MA, Wibberley S, Snell N, Rose DM, Swigris JJ. The need for patient-centred clinical research in idiopathic pulmonary fibrosis. BMC Med. 2015;13:240. https://doi.org/10.1186/s12916-015-0475-4.
7. Jones PW, Quirk FH, Baveystock CM, Littlejohns P. A self-complete measure of health status for chronic airflow limitation. The St. George's Respiratory Questionnaire. Am Rev Respir Dis. 1992;145:1321–7. https://doi.org/10.1164/ajrccm/145.6.1321.
8. Swigris JJ, Esser D, Conoscenti CS, Brown KK. 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. https://doi.org/10.1186/s12955-014-0124-1.
9. Yorke J, Jones PW, Swigris JJ. Development and validity testing of an IPF-specific version of the St George's Respiratory Questionnaire. Thorax. 2010;65:921–6.
10. Prior TS, Hoyer N, Hilberg O, Shaker SB, Davidsen JR, Bendstrup E. Responsiveness and minimal clinically important difference of SGRQ-I and K-BILD in idiopathic pulmonary fibrosis. Respir Res. 2020;21:91. https://doi.org/10.1186/s12931-020-01359-3.
11. Prior TS, Hoyer N, Hilberg O, Shaker SB, Davidsen JR, Rasmussen F, et al. Clusters of comorbidities in idiopathic pulmonary fibrosis. Respir Med. 2021;185:106490.
12. Patel AS, Siegert RJ, Brignall K, Gordon P, Steer S, Desai SR, et al. The development and validation of the King's Brief Interstitial Lung Disease (K-BILD) health status questionnaire. Thorax. 2012;67:804–10.
13. Eakin EG, Resnikoff PM, Prewitt LM, Ries AL, Kaplan RM. Validation of a new dyspnea measure: the UCSD Shortness of Breath Questionnaire. Chest. 1998;113:619–24.
14. Swigris JJ, Han M, Vij R, Noth I, Eisenstein EL, Anstrom KJ, et al. The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. Respir Med. 2012;106:1447–55.
15. Ware JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care. 1992;30:473–83.
16. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. J Man Manip Ther. 2009;17:163–70.
17. Lancaster LH. Utility of the six-minute walk test in patients with idiopathic pulmonary fibrosis. Multidisc Respir Med. 2018;13:45.
18. Raghu G, Collard HR, Egan JJ, Martinez FJ, Behr J, Brown KK, 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.
19. du Bois RM, Weycker D, Albera C, Bradford WZ, Costabel U, Kartashov A, 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;184:1382–9.
20. Patel AS, Siegert RJ, Keir GJ, Bajwah S, Barker RD, Maher TM, 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;107:1438–43.
21. Hathaway EH, Tashkin DP, Simmons MS. Intraindividual variability in serial measurements of DLCO and alveolar volume over one year in eight healthy subjects using three independent measuring systems. Am Rev Respir Dis. 1989;140:1818–22. https://doi.org/10.1164/ajrccm/140.6.1818.
22. Swigris JJ, Wamboldt FS, Behr J, du Bois RM, King TE, Raghu G, et al. The 6 minute walk in idiopathic pulmonary fibrosis: longitudinal changes and minimum important difference. Thorax. 2010;65:173–7.
23. Ley B, Ryerson CJ, Vittinghoff E, Ryu JH, Tomassetti S, Lee JS, et al. A multidimensional index and staging system for idiopathic pulmonary fibrosis. Ann Intern Med. 2012;156:684–91.
24. Terwee CB, Bot SDM, de Boer MR, van der Windt DAWM, Knol DL, Dekker J, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60:34–42.
25. Swigris JJ, Wilson H, Esser D, Conoscenti CS, Stansen W, Kline Leidy N, 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.
26. Revicki DA, Cella D, Hays RD, Sloan JA, Lenderking WR, Aaronson NK. Responsiveness and minimal important differences for patient reported outcomes. Health Qual Life Outcomes. 2006;4:70.
27. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol. 2003;56:395–407.
28. de Vet HCW, Ostelo RWJG, Terwee CB, van der Roer N, Knol DL, Beckerman H, et al. Minimally important change determined by a visual method integrating an anchor-based and a distribution-based approach. Qual Life Res. 2007;16:131–42. https://doi.org/10.1007/s11136-006-9109-9.
29. Swigris JJ, Brown KK, Behr J, du Bois RM, King TE, Raghu G, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. Respir Med. 2010;104:296–304. https://doi.org/10.1016/j.rmed.2009.09.006.
30. 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.
31. Furukawa T, Taniguchi H, Ando M, Kondoh Y, Kataoka K, Nishiyama O, et al. The St. George's Respiratory Questionnaire as a prognostic factor in IPF. Respir Res. 2017;18:18.
32. Case AH, Hellkamp AS, Neely ML, Bender S, Dilling DF, Gulati M, et al. Associations between patient-reported outcomes and death or lung transplant in idiopathic pulmonary fibrosis data from the idiopathic

Prior *et al. Respir Res*   (2021) 22:259

Page 10 of 10

pulmonary fibrosis prospective outcomes registry. Ann Am Thorac Soc. 2020;17:699–705.

33. Glaspole IN, Chapman SA, Cooper WA, Ellis SJ, Goh NS, Hopkins PM, et al. Health-related quality of life in idiopathic pulmonary fibrosis: data from the Australian IPF Registry. Respirology. 2017;22:950–6.

34. Nishiyama O, Taniguchi H, Kondoh Y, Kimura T, Kataoka K, Nishimura K, et al. Health-related quality of life does not predict survival in idiopathic pulmonary fibrosis. Sarcoidosis Vasc Diffuse Lung Dis. 2012;29:113–8.

35. Kreuter M, Swigris J, Pittrow D, Geier S, Klotsche J, Prasse A, et al. The clinical course of idiopathic pulmonary fibrosis and its association to quality of life over time: longitudinal data from the INSIGHTS-IPF registry. Respir Res. 2019;20:59.

36. Rönmark E, Lundqvist A, Lundbäck B, Nyström L. Non-responders to a postal questionnaire on respiratory symptoms and diseases. Eur J Epidemiol. 1999;15:293–9.

Publisher's Note