Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

.

_____ _____

Mohleen Kang Date

Approval Sheet

Assessing Responsiveness and Determining Minimal Clinically Important Differences In Idiopathic Pulmonary Fibrosis

By

Mohleen Kang
Master of Science
Clinical Research

_____

Greg S Martin, MD, MSc
Advisor

_____

Jordan Kempker, MD, MSc
Advisor

_____

Srihari Veeraraghavan, MD
Advisor

_____

Sushma Cribbs, MD, MSc
Committee Member

_____

Matthew Magee, PhD
Committee Member

Accepted:

_____

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

_____

Date

Assessing Responsiveness and Determining Minimal Clinically Important Differences In Idiopathic Pulmonary Fibrosis

By

Mohleen Kang
M.D., Virginia Commonwealth University School of Medicine, 2013
B.S., Virginia Commonwealth University, 2008

Advisors:
Greg S Martin, MD, MSc
Jordan A Kempker, MD, MSc
Srihari Veeraraghavan, MD

An abstract of
A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Master of Science
in Clinical Research
2021

ABSTRACT

Assessing Responsiveness and Determining Minimal Clinically Important Differences In Idiopathic Pulmonary Fibrosis
By Mohleen Kang

Introduction: Idiopathic pulmonary fibrosis (IPF) is a progressive lung disease with a median survival of 2-3 years after diagnosis. Current medications have been shown to decrease the decline in lung function but have no impact on patient-reported outcome measures (PROMs). The aim of this study was to assess the responsiveness of various physiologic measures and PROMs and to estimate Minimal Clinically Important Differences (MCID) values for worsening using anchor based methods.
Methods: We conducted secondary analyses of three randomized controlled trials (STEP-IPF, ACE-IPF and PANTHER-IPF) with different inclusion criteria and follow-up intervals. The Health Transition question in the Short Form Health Survey 36 (SF36) questionnaire was used as the anchor. Receiver operating curve analysis was used to assess responsiveness between the anchor and ten variables of interest (four physiologic measures and six PROMs) and area under the curve ≥ 0.70 was set as the threshold. To determine the MCID values, we used two anchor-based methods, one proposed by Jaeschke and another by Redelmeier.
Results: Only four variables met the responsiveness criteria: 1) Six-minute walk distance (6MWD), 2) St. George's Respiratory Questionnaire (SGRQ), 3) physical component score of SF36 (SF36 PCS), and 4) University of California, San Diego, Shortness of Breath Questionnaire (UCSD SOBQ). The MCID values for 6MWD were -75 meters and -43 meters over 24 weeks using Jaeschke and Redelmeier methods respectively. The MCID values for SF36 PCS over 60 weeks were -7 using both Jaeschke and Redelmeier methods. MCID values for SGRQ over 60 weeks were 11 and 10 using Jaeschke and Redelmeier methods respectively. MCID values for the UCSD SOBQ over 60 weeks were 11 using both Jaeschke and Redelmeier methods.
Conclusions: The MCID estimates of 6MWD, SGRQ, SF36, UCSD SOBQ varied considerably from previously proposed values. A single MCID value may not be applicable across all classes of disease severity or durations of follow-up time.

Assessing Responsiveness and Determining Minimal Clinically Important Differences In Idiopathic Pulmonary Fibrosis

By

Mohleen Kang
M.D., Virginia Commonwealth University School of Medicine, 2013
B.S., Virginia Commonwealth University, 2008

Advisors:
Greg S Martin, MD, MSc
Jordan A Kempker, MD, MSc
Srihari Veeraraghavan, MD

A thesis submitted to the Faculty of the James T. Laney School of Graduate Studies of Emory University in partial fulfillment of the requirements for the degree of Master of Science
in Clinical Research
2021

ACKNOWLEDGMENTS

This project would not have been possible without the support and guidance of several people. I would first like to thank Dr. David Guidot, Director of the Division of Pulmonary, Allergy, Critical Care and Sleep Medicine, for enthusiastically supporting my career goals and offering his wise counsel in times of uncertainty. I am extremely grateful to my senior mentor, Dr. Greg Martin, for overseeing my mentorship team and successfully guiding me throughout my master's degree. I want to also thank Dr. Srihari Veeraraghavan for inspiring my interest in interstitial lung diseases and helping formulate important research questions. I am deeply indebted to Dr. Jordan Kempker for spending countless hours helping me design and analyze my research project, editing all my writings, patiently listening to all my ideas, and most importantly for his profound belief in my abilities. I would also like to acknowledge Dr. Sushma Cribbs for her insightful review of my thesis. Lastly, I would like to thank faculty, staff and students of the Master of Science in Clinical Research program, for everything that I have learned in the last two years.

This Manuscript was prepared using ACE, PANTHER, STEPIPF Research Materials obtained from the NHLBI Biologic Specimen and Data Repository Information Coordinating Center and does not necessarily reflect the opinions or views of the ACE, PANTHER, STEPIPF or the NHLBI. This work was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR002378. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

TABLE OF CONTENTS

Page

INTRODUCTION

Idiopathic Pulmonary Fibrosis (IPF) is a chronic fibrosing lung disease that is progressive and has a median survival of 2-3 years after diagnosis (1). The disease progression is associated with increased symptom burden and is punctuated by episodic acute exacerbations that can lead to hospitalization and acute respiratory failure. There are currently two pharmacologic treatment options, Pirfenidone and Nintedanib, which have been shown to decrease the rate of annual decline of forced vital capacity (FVC) (2-5). Neither of these medications, however, has shown a difference in patient-reported outcomes measures (PROMs) as measured by the St. George's Respiratory Questionnaire (SGRQ) or the University of California, San Diego Shortness of Breath Questionnaire (UCSD SOBQ). This raises an important issue as to what minimal clinically important difference (MCID) in outcome measures such as FVC would be associated with clinically meaningful change in patients.

MCID is a threshold value for a change in a measure considered meaningful by the patient. It is a complementary approach to the practice of solely relying on statistical significance to determine important differences in a measure. Traditionally, MCID values have been estimated by triangulation of three methods: distribution-based, anchor-based, and expert opinion. Distribution-based methods do not incorporate patient input and, therefore, may not necessarily reflect patient-centered differences (6, 7). Current consensus approaches support using anchor-based over distribution-based methods (8, 9).

The overall aim of this exploratory study was to assess the responsiveness of various physiologic and PROMs and estimate MCID values using anchor-based methods by conducting secondary data analysis of three different IPF randomized controlled trials.

We used receiver operating curve analysis to assess responsiveness between the anchor and ten variables of interest (four physiologic measures and six PROMs). To determine the MCID values of variables that met responsiveness criteria (area under the curve ≥ 0.70), we used two anchor-based methods, one proposed by Jaeschke and another by Redelmeier (10-14).

The Health Transition question (SF 2) in the Short Form Health Survey 36 questionnaire was selected as the anchor. We used four available physiologic measurements: FVC, total lung capacity (TLC), diffusing capacity of the lung for carbon monoxide (DLCO) and 6-minute walk distance (6MWD). We used the following well-established PROMs: (SGRQ, Short Form Health Survey 36 (SF36) and UCSD SOBQ, EuroQoL questionnaire, Borg Dyspnea scale, and Investigating Choice Experiences for the Preferences of Older People Capability Instruments for Adults (ICECAP) questionnaire.

BACKGROUND

IPF is a chronic fibrosing lung disease that is progressive and has a very poor prognosis. As the name implies, the exact cause of the disease process is unknown. Patients often present with shortness of breath with exertion and cough, have characteristic findings on high-resolution computed tomography (HRCT) of lung fibrosis, and a restrictive pattern on pulmonary function tests. Diagnosis of IPF is a complex interpretation of historical, radiographic, functional, and sometimes histopathological components and requires an extensive exclusion of the other causes of interstitial lung diseases (15). Progression of IPF is associated with increased fibrosis as seen on HRCT scans of the lung, worsening pulmonary function tests leading to increased symptom burden, disability often requiring oxygen therapy and ultimately death. The median survival after diagnosis is 2-3 years and the progressive disease course is punctuated by episodic acute exacerbations that can lead to hospitalization and acute respiratory failure (1). Acute exacerbations are clinically significant events in themselves, and precede almost 46% of deaths in IPF and the median survival of patients after an acute exacerbation is approximately 3-4 months (16). Therefore, IPF is a fatal disease with significant degree of suffering for those afflicted.

There are currently two oral medications that are treatment options, Pirfenidone and Nintedanib, which have been shown to decrease the rate of decline of FVC, one of the commonly used physiologic measures of lung function (2). These medications do not cure IPF or reverse the fibrosis that is already present. While individual trials of Pirfenidone did not show any difference in mortality between treatment and placebo groups, a pooled analysis of the studies did show decreased all-cause mortality favoring pirfenidone (17, 18). A pooled analysis of Nintedanib trials showed a hazard ratio for time to on-treatment

mortality of 0.57 (95% CI 0.34-0.97) favoring Nintedanib but the hazard ratio for time to all-cause mortality was 0.70 (95%CI 0.46-1.08) (19). Nintedanib has also been shown to have benefit in time to first acute exacerbation compared to placebo (HR 0.53 95% CI 0.34-0.83) in the pooled analysis but there is lack of evidence of similar benefit from Pirfenidone (19, 20). These medications are also expensive at almost $100,000 per patient per year and have significant side effects such as diarrhea and photosensitive rash (21). Pooled analysis showed that almost 20.6% of patients on Nintedanib compared to 15% on placebo and 11.9% of patients on Pirfenidone compared to 8.7% on placebo discontinued the drug due to adverse events (18, 19). These results raise the important question as to whether these therapies are truly affecting patient-centered outcomes to justify their costs and side effects.

As far as PROMs are concerned, there was no significant difference seen in UCSD SOBQ at 52 weeks between the Pirfenidone and placebo groups in the ASCEND and CAPACITY trials (3, 5). A pooled analysis of the TOMORROW and INPULSIS studies did show that Nintedanib had lower adjusted change in total SGRQ score from baseline at 52 weeks compared to placebo (2.92 vs. 4.97 with p= 0.0095) (19). However, MCID of SGRQ in IPF patients has been previously estimated to be 5-8 points and, therefore, this difference may not be clinically important to patients (22). This raises the issue as to what magnitude of change in physiologic parameters such as FVC would be associated with the expected changes in PROMs such as UCSD SOBQ or SGRQ.

MCID is a threshold value for a change in a parameter that is considered meaningful by the patient rather than solely relying on statistically significant change in the parameter that may not be impactful to a patient's quality of life. Establishing MCID values for commonly used physiologic measures and PROMs can help research trials establish

endpoints that are not just statistically significant but also meaningful to the patients they are treating. MCID can be estimated by a triangulation of three approaches: 1) distribution method, 2) anchor method, and 3) expert opinion (23, 24). Anchor-based methods determine the MCID by associating the change in the numerical scale for a parameter to another independent assessment of improvement or worsening. Distribution-based methods, on the other hand, rely on the statistical properties of the distribution of a parameter to determine what magnitude of change is required to show that the change in a parameter is more than would be expected from chance alone. While there is no gold standard methodology to determine MCID values, some tools and consensus approaches have been proposed in the literature in an effort to standardize MCID estimation and they support using anchor-based over distribution-based methods (8, 9).

There has been some initial, focused investigation in the area of estimating MCID values for physiologic parameters and PROMs in IPF; however, these studies have some significant limitations. Among the ten articles that have studied MCID values of various measures in IPF (22, 25-33). Nine out of the ten studies utilized distribution-based methods to calculate MCID (22, 25-29, 31-33). Additionally, while these studies also used anchor-based methods along with distribution-based methods, some of the studies used mortality and or hospitalization as anchors, which while clinically important to patients, may determine "maximal" rather than "minimal" important changes (26-28). Similarly, physiologic measures, such as FVC, do not incorporate patient-centered input about change and may be less than ideal especially when used as sole anchors in a study (22, 25, 30-32). None of the studies used expert opinion or consensus based approaches to determine MCID.

METHODS

Hypothesis and Aim

We hypothesized that for a chronic progressive lung disease like IPF, most patients would either be unchanged or worsened clinically at the end of the specified follow-up period. Therefore, the aim of our study was to assess responsiveness of various physiologic measures and PROMs and estimate MCID values associated with patient worsening only.

Study Design

We conducted secondary analyses of data from three randomized controlled trials: Sildenafil Trial of Exercise Performance in Idiopathic Pulmonary Fibrosis (STEP-IPF), AntiCoagulant Effectiveness in Idiopathic Pulmonary Fibrosis (ACE-IPF), and Prednisone, Azathioprine, and N-Acetylcysteine: A Study That Evaluates Response in Idiopathic Pulmonary Fibrosis (PANTHER-IPF) (34-36). Data from these trials was obtained from the National Heart, Lung and Blood Institute (NHLBI) via the Biologic Specimen and Data Repository Information Coordinating Center (BioLINCC) program.

Study Characteristics

While all of these trials enrolled patients with IPF, each had different inclusion criteria and study durations (Table 1).

The STEP-IPF trial was a double blind, randomized, placebo-controlled trial of sildenafil in 180 patients, conducted at 14 US centers over two 12-week time periods from 2007-2009 (31). During the first 12 weeks, patients were randomized 1:1 to either placebo or sildenafil and the next 12-week period was an open label extension of sildenafil to both groups (31). The trial had the following inclusion criteria: 1) definite IPF as defined by a multidisciplinary approach at each center and 2) DLCO <35% (31). The major exclusion

criteria were: 1) 6MWD <50 m; 2) on nitrate therapy; 3) on treatment for pulmonary hypertension with prostaglandins, endothelin receptor antagonists, phosphodiesterase type 5 inhibitors; 4) resting oxygen saturation <92% on 6 liters per minute of supplemental oxygen; or 5) actively listed on lung transplantation (31). The study found no significant difference in 6MWD improvement of at least 20% from baseline between two groups (31).

The ACE-IPF was a double-blind, randomized, placebo-controlled trial of warfarin in 145 patients over 48 weeks in 22 US centers from 2009-2011 with the following inclusion criteria: 1) diagnosis of IPF according to modified American Thoracic Society (ATS) guidelines; 2) age 35-80 years; and 3) either worsening dyspnea or progressive fibrosis on radiologic imaging or one of the following: absolute decline of FVC ≥10%, decline in DLCO ≥15% or reduction in arterial oxygen saturation ≥5% (30). The trial excluded patients with the following: 1) current other medical indication for warfarin use; 2) increased risk of bleeding; or 3) actively listed for lung transplantation (30). The study was terminated earlier than a planned time period of 144 weeks after an interim analysis showed that patients on warfarin had higher mortality, hospitalization and severe side effects (29).

The PANTHER-IPF trial was a double blind, randomized, placebo-controlled trial of combinations of oral therapies in 155 patients with mild-moderate lung function impairment over 60 weeks at 25 US centers from 2009-2011 (30). The major inclusion criteria was: 1) diagnosis of IPF according to ATS criteria(37); 2) age 35-85 years; 3) FVC ≥50%; and 4) DLCO ≥30% (30). It excluded patients with the following: 1) evidence of emphysema greater than fibrosis from HRCT or pulmonary function tests; 2) actively listed for lung transplantation; 3) homozygous for low thiopurine-S-methyl transferase (TPMT)

levels; or 4) uncontrolled depression (30). The trial was stopped early after interim analysis found that the treatment group that received prednisone, azathioprine and n-acetylcysteine had increased mortality, hospitalizations and adverse events (30).

Given that the three studies had different inclusion and exclusion criteria and different follow-up time periods, three separate analyses following the same procedures were conducted for each. The time period for STEP-IPF cohort was set at 24 weeks, ACE-IPF was set at 48 weeks and PANTHER-IPF was set at 60 weeks. We used both the placebo and treatment arm participants in our analysis.

Measurements

For our anchor, we selected the Health Transition question (SF2) in the 36-Item Short Form Survey. SF2 asks the patients to rate their health on a five point Likert scale in response to the following question: "Compared with one year ago, how would you rate your health in general now?" Possible responses to this question were as follows: (1) "much better," (2) "somewhat better," (3) "same," (4) "somewhat worse," and (5) "much worse" (38). We analyzed patients with complete SF2 data at the end of the respective study follow-up time period.

The physiologic measures included in our analysis were FVC, TLC, DLCO, 6MWD. We evaluated both absolute change in percent predicted FVC and FVC in liters (L) separately. We also analyzed relative change in FVC in L which was expressed as a percentage. For DLCO we evaluated absolute difference in percent predicted DLCO and DLCO measured as ml/min/mmHg. The STEP-IPF dataset obtained from BioLINCC did not include percent predicted values for FVC and DLCO. We used NHANES spirometry reference values to compute percent predicted values for FVC for the STEP-IPF cohort

(39). Percent predicted values for DLCO were not computed for STEP-IPF cohort. For TLC, the absolute difference in TLC in liters was analyzed in ACE-IPF and PANTHER-IPF cohorts. The TLC values were not available in the STEP-IPF dataset. For 6MWD, we analyzed absolute difference in 6MWD in meters.

The PROMs we examined included Borg dyspnea scale, SF36 physical and mental component scores, EuroQol score index and visual analogue scores, SGRQ, UCSD SOBQ and ICECAP questionnaire. The STEP-IPF data set did not include total scores for SGRQ, SF36 physical and mental components, UCSD SOBQ, EuroQoL index and visual analogue scale or ICECAP questionnaire. We calculated the total scores for UCSD SOBQ and the EuroQol index and visual analogue scale (using the SAS code provided by EuroQol Group). We were unable to compute total scores for SGRQ, SF36 and ICECAP in the STEP-IPF cohort due to missing components.

Analytical Plan

All analyses were conducted using SAS version 9.4 (SAS Institute, Cary, NC) and IBM SPSS Statistics version 26 (SPSS Inc., Chicago, IL). We initially performed descriptive univariate analyses for each patient measure retaining all outliers in the analysis. We calculated mean change (score difference between follow-up and baseline) of each measure for patients in each of the categories in the SF2 question.

For MCID calculation we followed a specific, step-wise criteria for selecting variables for MCID estimation. Specifically, we conducted receiver operating curve analysis to assess responsiveness of the change in variable of interest with the anchor SF2 (dichotomous variable “same” and “somewhat worse”). Measures with both an area under the curve (AUC) ≥ 0.70 and with appropriate direction of response i.e. worsening scores

with worsening response to SF2 question were selected for further MCID determination. Given that there is no gold standard method for calculating anchor-based MCID, in our exploratory approach we utilized two published methods. First, we calculated the score difference of the measure from baseline to follow-up in patients who answered “somewhat worse” in response to SF2 as MCID (proposed by Jaschke) (10). Secondly, we calculated the mean change between the “same” and “somewhat worse” SF2 groups as MCID (proposed by Redelmeier) (11). All analyses were conducted using observed cases. If patients had missing data at follow-up, then those patients were not included in the MCID analysis.

RESULTS

Baseline Characteristics

A total of 140 patients had follow-up data at 24 weeks in the STEP-IPF cohort, 111 patients had follow-up data at 48 weeks in the ACE-IPF cohort and 228 patients had follow-up data at 60 weeks in the PANTHER-IPF trial. Participants from all three cohorts were predominantly male (71-81%) and white (92-96%). The STEP-IPF cohort had a mean (SD) age of 68.47 (9.11) years with mean (SD) percent predicted FVC of 58.52 (15.50)% and mean (SD) DLCO of 7.92 (2.12) ml/min/mmHg (Table 2). The ACE-IPF cohort had a mean (SD) age of 66.65 (7.49) years with a mean (SD) percent predicted FVC of 61.94 (15.19)% and mean (SD) DLCO of 36.16 (12.90) % (Table 3). The PANTHER-IPF cohort had mean (SD) age of 67.05 (8.32) years with a mean (SD) percent predicted FVC of 73.81(15.05)% and DLCO of 46.18 (11.36) % (Table 4).

Response to Anchor SF2

In the STEP-IPF cohort, 110 out of the 140 patients (78.6%) were either in the "same" or in the "somewhat worse" category according to SF2 response at follow-up (Table 5). 6MWD was the only measure in the STEP-IPF cohort that met the responsiveness criteria (AUC $\geq$ 0.70) for MCID estimation (Table 5). The AUC for other physiologic measures and PROMs in the STEP-IPF ranged from 0.55 - 0.68 (Table 5). In the ACE-IPF cohort, 98 out of the 111 patients (88.3%) with follow-up data at 48 weeks answered "same" or "somewhat worse" in response to the SF2 question at 48 weeks (Table 6). None of the physiologic measures or the PROMs in the ACE-IPF cohort met the prespecified responsiveness criteria for MCID determination with AUC ranging from 0.53 to 0.61 (Table 6). In the PANTHER-IPF cohort, 175 out of 228 patients (76.8%) answered

about the same or somewhat worse in response to the SF2 question (Table 7). In the PANTHER-IPF cohort, the physical component score of the SF36 questionnaire, the total SGRQ and UCSD SOBQ scores were the only measures that met criteria for next stage of MCID calculation (Table 7). The AUC for other physiologic measures and PROMs in the PANTHER-IPF ranged from 0.47 - 0.69 (Table 7).

MCID Values for Worsening

MCID values for 6MWD were -74.89 (95% CI -93.11, -56.66) and -42.59 (95% CI -75.95, -9.24) over 24 weeks using Jaeschke and Redelmeier methods respectively (Table 8). The MCID values for physical component score of SF36 over 60 weeks were -6.79 (95% CI -8.66, -4.92) using Jaeschke method and -6.73 (95% CI -8.91, -4.55) using Redelmeier method (Table 8). MCID values for total SGRQ score over 60 weeks were 10.95 (95% CI 7.81, 14.1) and 9.61 (95% CI 5.96, 13.25) using Jaeschke and Redelmeier methods respectively (Table 8). MCID values for the total UCSD SOBQ score over 60 weeks were 11.38 (95% CI 7.83, 14.93) and 10.78 (95% CI 6.46, 15.11) using Jaeschke and Redelmeier methods, respectively (Table 8).

DISCUSSION

This is the first study in IPF to conduct a comprehensive exploratory analysis of multiple physiologic measures and PROMs in three different cohorts using an anchor-based approach consistent with recently proposed standards in the MCID literature. This study demonstrates several key points: first, the MCID estimates of 6MWD, SGRQ, SF36, UCSD SOBQ were higher than previously calculated point estimates. These previous studies not only used different methodology, but in most instances, conducted their analyses on patients with different baseline disease severity and with different follow-up intervals which makes direct comparison difficult. Second, in our analysis, no one measure met responsiveness criteria in more than one cohort. Third, the variable FVC, the primary end point in major trials, did not meet responsiveness criteria in any of the three cohorts. This variation in responsiveness of outcome measures may be due to random chance, different duration of follow-up compared to the anchor, study procedures, or bias; or some combination of them all. Our findings demonstrate the complexities of MCID calculation which has large implications for trial design and evaluation.

Our study's results must be understood in the context of its limitations. First, we used a single anchor for our analysis. The Health Transition question (SF2) has been used in MCID determination in other studies and meets the requirements of patient reported anchor as first proposed by Jaeschke (10, 27). Others have argued that anchors with a single item are less reliable compared to multi-item anchors and therefore the results of our analysis should be confirmed with other anchors (6). Second, SF2 asks patient to recall their general health over the last one year which makes it prone to recall bias and using this to anchor changes over other time periods is not ideal. Third, the anchor-based methods

proposed by Jaeschke and Redelmeier used in our analysis are prone to regression to the mean phenomenon (12, 40). Finally, our study assessed responsiveness and estimated MCID but did not assess the validity or psychometric properties of these measures. Previous studies have evaluated convergent validity and some psychometric properties of 6MWD, SF36, SGRQ and UCSD SOBQ in IPF (22, 25, 32, 33, 41, 42). Even with these limitations, the MCID values estimated in our analysis represent some significant methodological strengths over prior IPF work.

Most MCID studies propose using a correlation coefficient (usually ≥ 0.3 or 0.5) to assess responsiveness of the change in the measure with the anchor (12, 40). This approach is suitable for diseases such as chronic pain where patients are expected to be categorized somewhat evenly into the five-point Likert scale categories of an anchor like SF2. However, in a chronic progressive disease like IPF, most patients fall into only two of the five anchor categories and using correlation coefficient may not accurately identify variables that are responsive to the anchor. Given the imbalance in categories, which was seen in all three cohorts in our study, the receiver operating curve analysis with AUC ≥ 0.70 was used to assess responsiveness of variables to a dichotomous anchor (14, 43).

Compared to previous MCID studies in IPF, we did not use distribution-based methods in our calculation. Distribution-based methods do not take into account patient's report of their health; they essentially report the minimal detectable change (MDC). However, MDC and MCID are two different concepts as illustrated by de Vet and Terwee (7). Previous MCID studies in IPF have used distribution-based methods along with anchor-based methods and have reported lower point estimates when compared to our calculated values. The MCID estimates for 6MWD at 75m in our analysis is much higher

compared to previously reported values ranging from 21.7 - 45m (26, 28, 32). The estimate for SF36 PCS of 7 points is also higher when compared to previous values of 3 points and 5 points (22, 31). While the difference in baseline disease severity and follow-up intervals in some of the previous studies makes direct comparison difficult, in certain cases our MCID values fall within the reported ranges of previous studies even if they are higher than the point estimates. For instance, only one study thus far has determined MCID estimates of total UCSD SOBQ scores and used the STEP-IPF cohort for their analysis (33). They reported an MCID estimate of 8 points for both improvement and worsening with a range of 5-11 over 24 weeks using SGRQ's activity domain for anchor-based method along with distribution-based methods (33). The UCSD SOBQ score did not meet responsiveness criteria in our analysis of STEP-IPF cohort but our reported anchor-based MCID values for UCSD SOBQ at 11 points over a 60 week time period using mild to moderate disease patients of the PANTHER-IPF trial is close to the reported range of 5-11 in the previous study.

Similarly, an earlier study reported an MCID of SGRQ as 7 points with a range of 5-10 using both anchor-based and distribution-based methods in IPF patients with mild to moderate severity (22). In our analysis, we estimated higher MCID of SGRQ of 11 and 10 points over 60-week time period for worsening using a similar mild to moderate category of patients which again falls within or close to the range of the previous study but higher than the reported point estimate. However, another more recent study estimated MCID for SGRQ in IPF using mild to moderate severity patients over 52 weeks and proposed a threshold of 4-5 points for both improvement and worsening using both distribution and anchor-based methods and is much lower than our estimate (25). Further research is needed

to study the impact of MCID methodology, disease severity, follow-up interval on MCID estimation and there are efforts underway to study some of these relationships in other diseases such as Asthma (44). A study of MCID of three questionnaires including SGRQ in COPD patients found stable MCID values over different follow up intervals ranging from 3 weeks to 12 months (45). A large real world dataset of IPF patients, such as the newer patient registries, with patients of varying disease severity and multiple follow-up measurements at set intervals may be useful for standardized MCID research of physiologic measures and PROMs, provided they have appropriate anchors for MCID estimation (46, 47).

Conclusions:

Our study highlights the fact the anchor-based MCID estimates of 6MWD, SGRQ, SF36, UCSD SOBQ in our study were considerably higher when compared to point estimates from previously proposed values. Further research is needed to assess MCID values of various physiologic measures and PROMs in IPF using a more current and standardized approach in different patient cohorts over different time periods to better design and evaluate clinical trials. There is further need to establish MCID of newer physiologic measures such as home spirometry and actigraphy (48, 49). PROMs designed specifically for IPF patients are also needed to better capture the patient experience in clinical trials since PROMs like SGRQ were developed for patients with obstructive diseases. The newly proposed Living with Idiopathic Pulmonary Fibrosis (L-IPF) questionnaire is one such endeavor to better incorporate the patient experience (50). With these advances, future intervention trials in IPF may be better poised to accurately evaluate patient quality of life.

REFERENCES

1. Raghu G, Richeldi L. Current approaches to the management of idiopathic pulmonary fibrosis. *Respir Med* 2017;129:24-30.
2. Raghu G, Rochwerg B, Zhang Y, et al. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline: Treatment of Idiopathic Pulmonary Fibrosis. An Update of the 2011 Clinical Practice Guideline. *American journal of respiratory and critical care medicine* 2015;192(2):e3-19.
3. Noble PW, Albera C, Bradford WZ, et al. Pirfenidone in patients with idiopathic pulmonary fibrosis (CAPACITY): two randomised trials. *Lancet* 2011;377(9779):1760-9.
4. Richeldi L, du Bois RM, Raghu G, et al. Efficacy and safety of nintedanib in idiopathic pulmonary fibrosis. *N Engl J Med* 2014;370(22):2071-82.
5. King TE, Jr., 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(22):2083-92.
6. Terwee CB, Roorda LD, Dekker J, et al. Mind the MIC: large variation among populations and methods. *J Clin Epidemiol* 2010;63(5):524-34.
7. de Vet HC, Terwee CB. The minimal detectable change should not replace the minimal important difference. *J Clin Epidemiol* 2010;63(7):804-5; author reply 6.
8. Devji T, Carrasco-Labra A, Qasim A, et al. Evaluating the credibility of anchor based estimates of minimal important differences for patient reported outcomes: instrument development and reliability study. *BMJ* 2020;369:m1714.

9. Johnston BC, Ebrahim S, Carrasco-Labra A, et al. Minimally important difference estimates and methods: a protocol. *BMJ Open* 2015;5(10):e007953.

10. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989;10(4):407-15.

11. Redelmeier DA, Lorig K. Assessing the Clinical Importance of Symptomatic Improvements: An Illustration in Rheumatology. *Archives of Internal Medicine* 1993;153(11):1337-42.

12. Angst F, Aeschlimann A, Angst J. The minimal clinically important difference raised the significance of outcome effects above the statistical level, with methodological implications for future studies. *J Clin Epidemiol* 2017;82:128-36.

13. Husted JA, Cook RJ, Farewell VT, et al. Methods for assessing responsiveness: a critical review and recommendations. *Journal of Clinical Epidemiology* 2000;53(5):459-68.

14. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation* 2018;27(5):1147-57.

15. Raghu G, Remy-Jardin M, Myers JL, et al. Diagnosis of Idiopathic Pulmonary Fibrosis. An Official ATS/ERS/JRS/ALAT Clinical Practice Guideline. *American journal of respiratory and critical care medicine* 2018;198(5):e44-e68.

16. Collard HR, Ryerson CJ, Corte TJ, et al. Acute Exacerbation of Idiopathic Pulmonary Fibrosis. An International Working Group Report. *American journal of respiratory and critical care medicine* 2016;194(3):265-75.

17. Nathan SD, Albera C, Bradford WZ, et al. Effect of pirfenidone on mortality: pooled analyses and meta-analyses of clinical trials in idiopathic pulmonary fibrosis. *Lancet Respir Med* 2017;5(1):33-41.

18. Noble PW, Albera C, Bradford WZ, et al. Pirfenidone for idiopathic pulmonary fibrosis: analysis of pooled data from three multinational phase 3 trials. *The European respiratory journal* 2016;47(1):243-53.

19. Richeldi L, Cottin V, du Bois RM, et al. Nintedanib in patients with idiopathic pulmonary fibrosis: Combined evidence from the TOMORROW and INPULSIS((R)) trials. *Respir Med* 2016;113:74-9.

20. Torrisi SE, Pavone M, Vancheri A, et al. When to start and when to stop antifibrotic therapies. *European respiratory review : an official journal of the European Respiratory Society* 2017;26(145).

21. Owens GM. Strategies to manage costs in idiopathic pulmonary fibrosis. *The American journal of managed care* 2017;23(11 Suppl):S191-s6.

22. Swigris JJ, Brown KK, Behr J, et al. The SF-36 and SGRQ: validity and first look at minimum important differences in IPF. *Respiratory medicine* 2010;104(2):296-304.

23. McGlothlin AE, Lewis RJ. Minimal clinically important difference: defining what really matters to patients. *Jama* 2014;312(13):1342-3.

24. Jones PW, Beeh KM, Chapman KR, et al. Minimal clinically important differences in pharmacological trials. *American journal of respiratory and critical care medicine* 2014;189(3):250-5.

25. 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(1):e000278.

26. 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(9):1231-7.

27. 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. *American journal of respiratory and critical care medicine* 2011;184(12):1382-9.

28. Nathan SD, du Bois RM, Albera C, et al. Validation of test performance characteristics and minimal clinically important difference of the 6-minute walk test in patients with idiopathic pulmonary fibrosis. *Respir Med* 2015;109(7):914-22.

29. Nolan CM, Delogu V, Maddocks M, et al. Validity, responsiveness and minimum clinically important difference of the incremental shuttle walk in idiopathic pulmonary fibrosis: a prospective study. *Thorax* 2017.

30. 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(1):91.

31. Witt S, Krauss E, Barbero MAN, et al. Psychometric properties and minimal important differences of SF-36 in Idiopathic Pulmonary Fibrosis. *Respir Res* 2019;20(1):47.

32. Swigris JJ, Wamboldt FS, Behr J, et al. The 6 minute walk in idiopathic pulmonary fibrosis: longitudinal changes and minimum important difference. *Thorax* 2010;65(2):173-7.

33. Swigris JJ, Han M, Vij R, et al. The UCSD shortness of breath questionnaire has longitudinal construct validity in idiopathic pulmonary fibrosis. *Respiratory medicine* 2012;106(10):1447-55.

34. Noth I, Anstrom KJ, Calvert SB, et al. A placebo-controlled randomized trial of warfarin in idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* 2012;186(1):88-95.

35. Idiopathic Pulmonary Fibrosis Clinical Research N, Raghu G, Anstrom KJ, et al. Prednisone, azathioprine, and N-acetylcysteine for pulmonary fibrosis. *N Engl J Med* 2012;366(21):1968-77.

36. Idiopathic Pulmonary Fibrosis Clinical Research N, Zisman DA, Schwarz M, et al. A controlled trial of sildenafil in advanced idiopathic pulmonary fibrosis. *N Engl J Med* 2010;363(7):620-8.

37. American Thoracic Society. Idiopathic pulmonary fibrosis: diagnosis and treatment. International consensus statement. American Thoracic Society (ATS), and the European Respiratory Society (ERS). *American journal of respiratory and critical care medicine* 2000;161(2 Pt 1):646-64.

38. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992;30(6):473-83.

39. Hankinson JL, Odencrantz JR, Fedan KB. Spirometric reference values from a sample of the general US population. *American journal of respiratory and critical care medicine* 1999;159(1):179-87.

40. Malec JF, Ketchum JM. A Standard Method for Determining the Minimal Clinically Important Difference for Rehabilitation Measures. *Arch Phys Med Rehabil* 2020;101(6):1090-4.

41. Swigris JJ, Esser D, Wilson H, et al. Psychometric properties of the St George's Respiratory Questionnaire in patients with idiopathic pulmonary fibrosis. *Eur Respir J* 2017;49(1).

42. Swigris JJ, Streiner DL, Brown KK, et al. Assessing exertional dyspnea in patients with idiopathic pulmonary fibrosis. *Respiratory medicine* 2014;108(1):181-8.

43. De Vet HC, Terwee CB, Mokkink LB, et al. *Measurement in medicine: a practical guide*. Cambridge University Press; 2011.

44. Lanario J, Jones R, Hyland M, et al. Is the minimally clinically important difference (MCID) fit for purpose? A planned study using the SAQ. *European Respiratory Journal* 2020;56(suppl 64):2241.

45. Alma H, de Jong C, Jelusic D, et al. The minimal clinically important difference for COPD health status tools measured with global ratings of change during different time periods. *European Respiratory Journal* 2016;48(suppl 60):PA873.

46. Guenther A, Krauss E, Tello S, et al. The European IPF registry (eurIPFreg): baseline characteristics and survival of patients with idiopathic pulmonary fibrosis. *Respir Res* 2018;19(1):141.

47. Wang BR, Edwards R, Freiheit EA, et al. The Pulmonary Fibrosis Foundation Patient Registry: Rationale, Design, and Methods. *Ann Am Thorac Soc* 2020.

48. Russell AM, Adamali H, Molyneaux PL, et al. Daily Home Spirometry: An Effective Tool for Detecting Progression in Idiopathic Pulmonary Fibrosis. *Am J Respir Crit Care Med* 2016;194(8):989-97.

49. Nathan SD, Flaherty KR, Glassberg MK, et al. A Randomized, Double-Blind, Placebo-Controlled Study of Pulsed, Inhaled Nitric Oxide in Subjects at Risk of Pulmonary Hypertension Associated With Pulmonary Fibrosis. *Chest* 2020;158(2):637-45.

50. Swigris JJ, Andrae DA, Churney T, et al. Development and Initial Validation Analyses of the Living with Idiopathic Pulmonary Fibrosis (L-IPF) Questionnaire. *American Journal of Respiratory and Critical Care Medicine* 2020(ja).

TABLES

Table 1. Description of the three randomized control trials of idiopathic pulmonary fibrosis in adult patients used for secondary data analysis

Study Name	Enrollment	Number of Centers	Sample Size	Disease Severity/ Phenotype	Treatment	Time Period	Findings
STEP-IPF	2007-2009	14 US centers	180	Severe (DLCO <35%)	Sildenafil	Two 12 weeks: 1st placebo vs. sildenafil, 2nd sildenafil both groups	No significant improvement in 6MWD between groups
ACE-IPF	2009-2011	22 US centers	145	Progressive phenotype (either worsening dyspnea or absolute decline of FVC ≥10%, DLCO decline ≥15%, arterial oxygen saturation decline ≥5% or worsening radiographic findings)	Warfarin	48 weeks	Study terminated earlier since patients on warfarin had higher mortality, hospitalization and severe side effects
PANTHER-IPF	2009-2011	25 US centers	155	Mild to moderate (FVC ≥50% and DLCO ≥30%)	Prednisone,azathioprine and n-acetylcysteine	60 weeks	Trial stopped early. Treatment group with increased mortality, hospitalizations and adverse events

Abbreviations:

STEP-IPF Sildenafil Trial of Exercise Performance in Idiopathic Pulmonary Fibrosis
DLCO Diffusing capacity for carbon monoxide
6MWD Six minute walk distance
FVC Forced vital capacity
ACE-IPF AntiCoagulant Effectiveness in Idiopathic Pulmonary Fibrosis
PANTHER-IPF Prednisone, Azathioprine, and N-Acetylcysteine: A Study That Evaluates Response in Idiopathic Pulmonary Fibrosis

Table 2. Baseline characteristics of STEP-IPF patients with Health Transition question (SF2) data at 24 Weeks

Characteristic	Total Subjects (N=140)	Treatment (N=69)	Placebo (N=71)
Demographic Characteristics			
Age, years	68.47 (9.11)	69.66 (8.51)	67.35 (9.56)
Male, N (%)	114 (81.43)	58 (84.06)	56 (78.87)
White, N (%) †	130 (92.86)	63 (91.3%)	67 (94.36)
Hispanic or Latino, N (%)	10 (7.14)	6 (8.70)	4 (5.63)
Clinical Characteristics			
Past or Current Smoker, N (%)	112 (80.00)	54 (78.26)	58 (81.69)
Pulmonary Function Testing			
FEV1 %*	65.10 (16.23)	63.95 (16.88)	66.18 (15.65)
FEV1, L	1.95 (0.57)	1.91 (0.57)	1.98 (0.57)
FVC %*	58.52 (15.50)	56.37 (14.97)	60.51 (15.82)
FVC, L	2.37 (0.76)	2.29 (0.72)	2.45 (0.79)
TLC, L	3.68 (1.09)	3.62 (1.00)	3.74 (1.17)
DLCO, ml/min/mmHg	7.92 (2.12)	7.77 (1.94)	8.06 (2.28)
6MWD, meters	280.07 (112.27)	269.67 (99.88)	290.47 (123.28)
Quality of Life Scores			
Pre Borg Dyspnea Scale	0.79 (1.17)	0.84 (1.09)	0.74 (1.26)
Post Borg Dyspnea Scale	4.03 (8.09)	3.51 (1.68)	4.54 (11.25)
Total UCSD Score‡	46.82 (20.79)	49.61 (21.61)	44.21 (19.81)
EuroQol Visual Analog Score	69.34 (16.61)	68.41 (16.53)	70.26 (16.76)
Euroqol Index Score	0.80 (0.13)	0.79 (0.15)	0.80 (0.11)

*All values are percent predicted according to age, height, sex and race/ethnicity
† N Missing 10: 6 in treatment group and 4 in placebo
‡ N missing 22: 12 in treatment arm and 10 in placebo group
All other variables with N missing <5
FEV1 Forced expiratory volume in 1 second
FVC Forced vital capacity
TLC Total lung capacity
DLCO Diffusing capacity for carbon monoxide
6MWD Six-minute walk distance
SF36 36 Item Short Form Survey
SGRQ St. George's Respiratory Questionnaire
UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire
ICECAP Investigating Choice Experiences for the Preferences of Older People Capability Instruments for Adults

Table 3. Baseline characteristics of ACE-IPF patients with Health Transition question (SF2) data at 48 weeks

Characteristic	Total Subjects (N=111)	Treatment (N=54)	Placebo (N=57)
Demographic Characteristics			
Age in years, mean (SD)	66.65 (7.49)	66.65 (7.35)	66.64 (7.68)
Male, N (%)	79 (71.17)	35 (64.81)	44 (77.19)
White, N (%)	103 (92.79)	50 (92.59)	53 (92.98)
Minority, N (%)	14 (21.61)	7 (12.96)	7 (12.28)
Clinical Characteristics			
Past or Current Smoker, N (%)	83 (74.77)	37 (68.52)	46 (80.70)
Prednisone Treatment at Randomization, N (%)	27 (24.32)	10 (18.52)	17 (29.82)
Years since IPF diagnosis, mean (SD)	0.24 (0.43)	0.19 (0.39)	0.30 (0.46)
Pulmonary Function Testing, mean (SD)			
FVC, L	2.48 (0.78)	2.42 (0.77)	2.53 (0.79)
FVC %*	61.94 (15.19)	61.77 (15.94)	62.10 (14.58)
FEV1, L	2.03 (0.61)	2.00 (0.61)	2.06 (0.61)
FEV1 %*	65.68 (15.53)	66.28 (16.86)	65.12 (14.28)
DLCO, ml/min/mmHg	10.77 (4.41)	10.69 (4.03)	10.84 (4.77)
DLCO %*	36.16 (12.90)	36.39 (12.08)	35.94 (13.75)
TLC, L	3.75 (0.99)	3.71 (0.93)	3.78 (1.05)
TLC%*	59.38 (58.72)	59.40 (13.22)	59.36 (13.60)
6MWD, meters	297.90 (128.20)	303.73 (118.35)	292.59 (137.40)
Quality of Life Scores, mean (SD)			
Pre Borg Dyspnea Scale	0.59 (0.91)	0.48 (0.92)	0.69 (0.89)
Post Borg Dyspnea Scale	2.54 (1.54)	2.28 (1.53)	2.78 (1.53)
Total SGRQ Score	45.13 (16.19)	41.99 (16.08)	48.10 (15.85)
SF36 Physical Component Score	38.13 (8.66)	40.55 (8.37)	35.84 (8.36)
SF36 Mental Component Score	53.37 (8.33)	53.51 (7.37)	53.24 (9.21)
Total UCSD SOBQ Score	34.30 (20.71)	29.35 (17.61)	38.98 (22.42)
EuroQoL Index Score	0.76 (0.19)	0.78 (0.19)	0.75 (0.19)

EuroQol Visual Analog Score	74.90 (15.40)	75.72 (15.16)	74.14 (15.72)
ICECAP Score	0.86 (0.09)	0.87 (0.09)	0.85 (0.10)

*All values are percent predicted according to age, height, sex and race/ethnicity
FEV1 Forced expiratory volume in 1 second
FVC Forced vital capacity
TLC Total lung capacity
DLCO Diffusing capacity for carbon monoxide
6MWD Six-minute walk distance
SF36 36 Item Short Form Survey
SGRQ St. George's Respiratory Questionnaire
UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire
ICECAP Investigating Choice Experiences for the Preferences of Older People Capability Instruments for Adults

Table 4. Baseline characteristics of PANTHER IPF patients with Health Transition question (SF2) data at 60 weeks

Characteristic	**Total Subjects (N=228)**	**Treatment (N=110)**	**Placebo (N=118)**
Demographic Characteristics			
Age, years	67.05 (8.32)	67.50 (8.43)	66.63 (8.24)
Male, N (%)	173 (75.88)	87 (79.09)	86 (72.88)
White, N (%)	219 (96.05)	105 (95.45)	114 (96.61)
Minority, N (%)	15 (6.58)	7 (6.36)	8 (6.78)
Clinical Characteristics			
Past or Current Smoker, N (%)	164 (72.25)	78 (71.56)	86 (72.88)
Years since IPF diagnosis, median (IQR)	0.72 (0.29-1.54)	0.62 (0.29-1.54)	0.79 (0.28 -1.81)
Pulmonary Function Testing			
FVC, L	2.96 (0.80)	2.99 (0.82)	2.94 (0.78)
FVC % *	73.81 (15.05)	73.65 (15.70)	73.95 (14.49)
FEV1, L	2.44 (0.64)	2.44 (0.64)	2.44 (0.64)
FEV1 % *	78.85 (16.29)	78.06 (16.42)	79.59 (16.21)
DLCO, ml/min/mmHg	13.63 (3.76)	13.58 (3.69)	13.67 (3.85)
DLCO %*	46.18 (11.36)	45.70 (10.74)	46.63 (11.94)
TLC, L	4.38 (1.03)	4.45 (1.05)	4.33 (1.01)
6MWD, meters	383.71 (107.47)	385.34 (111.03)	382.21 (104.54)
Quality of Life Scores			
Pre Borg Dyspnea Scale	0.41 (0.81)	0.44 (0.88)	0.38 (0.74)
Post Borg Dyspnea Scale	2.41 (1.66)	2.27 (1.36)	2.53 (1.89)
Total SGRQ Score	38.21 (16.80)	38.72 (16.41)	37.74 (17.21)
SF36 Physical Component Score	41.18 (9.15)	41.47 (8.95)	40.91 (9.36)
SF36 Mental Component Score	54.65 (7.90)	53.87 (8.19)	55.38 (7.58)
Total UCSD Score	25.91 (17.54)	25.35 (16.80)	26.42 (18.26)
EuroQoL Index Score	0.83 (0.16)	0.82 (0.17)	0.84 (0.14)
EuroQol Visual Analog Score	78.25 (14.51)	78.38 (15.30)	78.13 (13.76)
ICECAP Score	0.88 (0.08)	0.88 (0.08)	0.88 (0.09)

*All values are percent predicted according to age, height, sex and race/ethnicity
FEV1 Forced expiratory volume in 1 second
FVC Forced vital capacity
TLC Total lung capacity
DLCO Diffusing capacity for carbon monoxide
6MWD Six-minute walk distance
SF36 36 Item Short Form Survey
SGRQ St. George's Respiratory Questionnaire
UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire
ICECAP Investigating Choice Experiences for the Preferences of Older People
Capability Instruments for Adults

Table 5. Change in physiologic and patient reported outcome measures over 24 weeks by Health Transition question (SF2) categorical responses in patients with idiopathic pulmonary fibrosis in the STEP-IPF trial

Variable	All Patients, Mean (SD) (N=140)	SF2 Response at 24-Week Follow-Up, Mean (SD)					AUC (95% CI)†
		Much Better (N= 7)	Somewhat Better (N=13)	About the Same (N= 48)	Somewhat worse (N=62)	Much Worse (N=10)	
Physiologic Measures							
ΔFVC%*	-2.72 (4.66)	-1.89 (4.49)	0.55(3.91)	-1.78 (4.52)	-3.68 (3.96)	-7.46 (6.75)	0.62 (0.51-0.72)
ΔFVC, L	-0.11 (0.20)	-0.09 (0.19)	0.04 (0.18)	-0.07 (0.19)	-0.15 (0.16)	-0.33 (0.25)	0.62 (0.51-0.72)
RFVC, %	-4.88 (8.38)	-3.44 (8.49)	1.13 (8.00)	-2.58 (7.50)	-6.98 (7.35)	-13.03 (10.67)	0.65 (0.55-0.76)
ΔDLCO, ml/min/mmHg	-0.48 (1.53)	-0.47 (1.54)	0.00 (1.30)	-0.16 (1.42)	-0.66 (0.51)	-1.65 (2.10)	0.55 (0.44-0.66)
Δ6MWD, meters	**-58.66 (96.44)**	**-3.71 (42.33)**	**-35.08 (144.51)**	**-32.29 (95.46)**	**-74.89 (69.32)**	**-163.00 (122.93)**	**0.72 (0.61-0.83)**
Subjective Measures							
ΔPre Borg Dyspnea Score	1.62 (11.65)	0.07 (0.93)	0.04 (0.78)	0.13 (1.01)	3.50 (17.27)	0.06 (3.17)	0.58 (0.47-0.69)
ΔPost Borg Dyspnea Score	1.06 (14.27)	-14.29 (35.62)	-1.00 (1.73)	4.10 (19.23)	0.76 (1.89)	2.00 (3.20)	0.59 (0.49-0.70)
ΔEuroQol Index Score	-0.05 (0.17)	0.09 (0.21)	-0.05 (0.20)	0.01 (0.12)	-0.08 (0.16)	-0.20 (0.21)	0.68 (0.58-0.78)
ΔEuroQol Visual Analogue Score	-1.32 (19.69)	26.57 (24.11)	8.62 (17.57)	1.52 (16.97)	-5.84 (16.57)	-21.89 (0.42)	0.66 (0.55-0.76)
ΔUCSD SOBQ Total Score‡	5.92 (16.6)	-8.00 (7.94)	-5.33 (13.66)	2.60 (9.92)	7.90 (19.10)	24.00 (13.83)	0.55 (0.43-0.68)

*All values are percent predicted according to age, height, sex and race/ethnicity

† Receiver operating curve comparing dichotomous SF2 response (about the "same" vs. "somewhat worse") and mean change score of variable

‡ N missing 39 (4 “much better”, 7 “somewhat better”, 13 “same”, 12 “somewhat worse”, 3 “much worse”). For all other variables the SF2 columns had N missing ≤ 5.
Δ Absolute change over 24 weeks
Abbreviations:
AUC Area under curve
FVC Forced vital capacity
RFVC Relative difference in forced vital capacity
TLC Total lung capacity
DLCO Diffusing capacity for carbon monoxide
6MWD Six-minute walk distance
UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire

Table 6. Change in physiologic and patient reported outcome measures over 48 weeks by Health Transition question (SF2) categorical responses in patients with idiopathic pulmonary fibrosis in the ACE-IPF trial

Variable	All Patients, Mean (SD), (N=111)	SF2 Response at 48-Week Follow-Up, Mean (SD)				AUC (95%CI) †
		Somewhat Better (N=6)	About the Same (N= 57)	Somewhat worse (N=41)	Much Worse (N=7)	
Physiologic Measures						
ΔFVC%*	-2.11 (6.61)	-1.00 (7.18)	-1.23 (5.61)	-1.73 (6.14)	-12.39 (8.94)	0.54 (0.41-0.66)
ΔFVC, L	-0.09 (0.26)	-0.04 (0.29)	-0.05 (0.20)	-0.08 (0.25)	-0.50 (0.40)	0.55 (0.42-0.69)
RFVC, %	-3.38 (10.61)	-0.21 (10.15)	-1.66 (8.33)	-3.23 (10.78)	-20.95 (12.53)	0.54 (0.42-0.66)
ΔTLC, L	0.00 (0.59)	0.15 (0.74)	0.06 (0.55)	-0.05 (0.60)	-0.34 (0.69)	0.58 (0.47-0.70)
ΔDLCO%*‡	-4.34 (9.80)	-0.23 (5.02)	-2.15 (6.38)	-7.22 (13.54)	-12.29 (7.13)	0.60 (0.48-0.73)
ΔDLCO, ml/min/mmHg ‡	-1.34 (2.94)	-0.23 (1.32)	-0.64 (1.82)	-2.23 (4.11)	-3.83 (2.31)	0.61 (0.48-0.74)
Δ6MWD, meters	-37.50 (114.96)	-18.81 (48.96)	-19.60 (119.23)	-49.31 (107.52)	-129.20 (119.90)	0.60 (0.48-0.71)
Subjective Measures						
ΔPre Borg Dyspnea Score	0.25 (1.29)	-0.30 (1.64)	0.01 (1.18)	0.57 (1.36)	1.10 (0.89)	0.60 (0.48-0.72)
ΔPost Borg Dyspnea Score§	0.49 (1.82)	-0.40 (1.34)	0.23 (1.60)	0.51 (1.55)	4.00 (2.65)	0.56 (0.44-0.69)
ΔSF36 Physical Component Score	-2.49 (7.15)	2.80 (9.21)	-1.60 (6.38)	-2.76 (6.73)	-12.77 (4.86)	0.55 (0.43-0.67)
ΔSF36 Mental Component Score	-1.56 (7.33)	1.81 (5.45)	-0.16 (6.65)	-2.81 (8.04)	-8.70 (4.35)	0.57 (0.45-0.68)
ΔEuroQol Index Score	-0.03 (0.18)	0.07 (0.08)	-0.01 (0.18)	-0.05 (0.20)	-0.07 (0.21)	0.54 (0.42-0.66)

ΔEuroQol Visual Analogue Score	-5.86 (15.19)	1.80 (11.01)	-3.21 (11.80)	-7.46 (17.35)	-23.57 (17.49)	0.59 (0.47-0.70)
ΔTotal SGRQ Score	4.54 (11.62)	0.57 (21.17)	2.41 (8.44)	5.20 (11.16)	20.73 (14.27)	0.58 (0.46-0.70)
ΔUCSD SOBQ Total Score	8.42 (16.65)	8.83 (12.25)	5.54 (10.86)	6.98 (16.65)	40.00(27.45)	0.53 (0.41-0.64)
ΔICECAP Score	-0.02 (0.10)	0.04 (0.12)	0.00(0.06)	-0.02 (0.11)	-0.16 (0.08)	0.59 (0.47-0.71)

*All values are percent predicted according to age, height, sex and race/ethnicity
† Receiver operating curve comparing dichotomous SF2 response (about the "same" vs. "somewhat worse") and mean change score of variable
‡ N missing 12 (0 "somewhat better", 3 "same", 8 "somewhat worse", 1 "much worse")
§ N missing 15 (1 "somewhat better", 6 "same", 6 "somewhat worse", 2 "much worse")
For all other variables the SF2 columns had N missing ≤ 5.
Δ Absolute change over 48 weeks
Abbreviations:
AUC Area under curve
FVC Forced vital capacity
RFVC Relative difference in forced vital capacity
TLC Total lung capacity
DLCO Diffusing capacity for carbon monoxide
6MWD Six-minute walk distance
SF36 36 Item Short Form Survey
SGRQ St. George's Respiratory Questionnaire
UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire
ICECAP Investigating Choice Experiences for the Preferences of Older People Capability Instruments for Adults

Table 7. Change in physiologic and patient reported outcome measures over 60 weeks by Health Transition question (SF2) categorical responses in patients with idiopathic pulmonary fibrosis in the PANTHER-IPF trial

Variable	All Patients, Mean (SD) (N=228)	SF2 Response at 60-Week Follow-Up, Mean (SD)					AUC (95% CI)†
		Much Better (N= 11)	Somewhat Better (N=32)	About the Same (N= 101)	Somewhat worse (N=74)	Much Worse (N=10)	
Physiologic Measures							
ΔFVC%*	-4.09 (6.75)	1.18 (7.31)	-2.64 (6.77)	-2.73 (5.80)	-6.65 (5.97)	-10.09 (11.13)	0.68 (0.60-0.76)
ΔFVC, L	-0.16 (0.28)	0.05 (0.29)	-0.10 (0.29)	-0.10 (0.23)	-0.27 (0.25)	-0.40 (0.44)	0.68 (0.60-0.76)
RFVC, %	-5.74 (9.99)	1.93 (10.81)	-3.46 (9.63)	-3.68 (8.06)	-9.46 (8.91)	-15.82 (19.43)	0.69 (0.61-0.77)
ΔTLC, L	-0.15 (0.49)	-0.01 (0.36)	0.07 (0.51)	-0.12 (0.44)	-0.27 (0.49)	-0.56 (0.63)	0.64 (0.55-0.73)
ΔDLCO%*	-4.28 (7.62)	0.38 (6.17)	-4.86 (5.91)	-3.32 (7.18)	-5.16 (8.06)	-14.78 (10.07)	0.59 (0.50-0.68)
ΔDLCO, ml/min/mmHg	-1.27 (2.21)	0.23 (1.72)	-1.37 (1.65)	-0.98 (1.99)	-1.58 (2.44)	-4.46 (2.91)	0.58 (0.50-0.67)
Δ6MWD, meters	-34.53 (100.53)	-2.32 (46.60)	-18.23 (119.36)	-22.64 (77.60)	-52.11 (113.71)	-229.00 (91.69)	0.57 (0.48-0.66)
Subjective Measures							
ΔPre Borg Dyspnea Score	0.29 (1.39)	-0.27 (0.52)	-0.39 (0.90)	0.39 (1.06)	0.51 (1.87)	1.30 (1.72)	0.51 (0.42-0.60)
ΔPost Borg Dyspnea Score	0.32 (0.92)	-1.27 (2.04)	-0.47 (1.78)	0.26 (1.61)	0.83 (2.01)	3.63 (1.60)	0.61 (0.52-0.69)
ΔSF36 Physical Component Score	**-2.25 (8.07)**	**8.00 (11.57)**	**-0.92 (6.26)**	**-0.06 (5.72)**	**-6.79 (8.07)**	**-5.29 (9.26)**	**0.75 (0.67-0.83)**
ΔSF36 Mental Component Score	-1.35 (8.61)	-0.29 (12.78)	-1.43 (9.66)	-0.41 (6.05)	-0.56 (7.61)	-17.35 (13.32)	0.47 (0.38-0.56)

ΔEuroQol Index Score	-0.04 (0.17)	-0.01 (0.19)	-0.06 (0.13)	0.01 (0.17)	-0.07 (0.14)	-0.29 (0.32)	0.66 (0.58-0.75)
ΔEuroQol Visual Analogue Score‡	-1.50 (17.19)	15.90 (16.39)	1.07 (10.75)	1.53 (16.35)	-5.79 (16.14)	-25.67 (18.86)	0.63 (0.54-0.72)
ΔTotal SGRQ Score§	**4.54 (13.74)**	**-13.70 (14.55)**	**-0.97 (9.60)**	**1.35 (9.18)**	**10.95 (13.19)**	**25.86 (17.46)**	**0.71 (0.63-0.79)**
ΔUCSD SOBQ Total Score	**5.72 (15.94)**	**-4.45 (11.79)**	**3.19 (9.60)**	**0.59 (12.81)**	**11.38 (15.31)**	**34.90 (23.65)**	**0.72 (0.65-0.80)**
ΔICECAP Score¶	-0.01 (0.09)	0.04 (0.08)	-0.01 (0.08)	0.01 (0.07)	-0.03 (0.09)	-0.16 (0.19)	0.69 (0.61-0.78)

*All values are percent predicted according to age, height, sex and race/ethnicity

† Receiver operating curve comparing dichotomous SF2 response (about the "same" vs. "somewhat worse") and mean change score of variable

‡ N missing 19 (1 "much better", 5 "somewhat better", 9 "same", 3 "somewhat worse", 1 "much worse")

§ N missing 13 (0 "much better", 2 "somewhat better", 8 "same", 3 "somewhat worse", 0 "much worse")

¶ N missing 12 (0 "much better", 0 "somewhat better", 6 "same", 6 "somewhat worse", 0 "much worse")

For all other variables the SF2 columns had N missing ≤ 5.

Δ Absolute change over 60 weeks

Abbreviations:

AUC Area under curve

FVC Forced vital capacity

RFVC Relative difference in forced vital capacity

TLC Total lung capacity

DLCO Diffusing capacity for carbon monoxide

6MWD Six-minute walk distance

SF36 36 Item Short Form Survey

SGRQ St. George's Respiratory Questionnaire

UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire

ICECAP Investigating Choice Experiences for the Preferences of Older People Capability Instruments for Adults

Table 8. Anchor-based estimates of Minimal Clinically Important Difference (MCID) for worsening in idiopathic pulmonary fibrosis

Variable	Time Period*	MCID Jaeschke† (95% CI)	MCID Redelmeier‡ (95% CI)
6MWD, m	24 Weeks	-74.89 (-93.11, -56.66)	-42.59 (-75.95, -9.24)
SF36 Physical Component Score	60 Weeks	-6.79 (-8.66, -4.92)	-6.73 (-8.91, -4.55)
Total SGRQ Score	60 Weeks	10.95 (7.81, 14.1)	9.61 (5.96, 13.25)
UCSD SOBQ Total Score	60 Weeks	11.38 (7.83, 14.93)	10.78 (6.46, 15.11)

* 24 Weeks from STEP-IPF cohort and 60 Weeks from PANTHER-IPF cohort
† Jaeschke: Score change in "somewhat worse" Health Transition question (SF2) group from baseline to follow-up
‡ Redelmeier: Difference between mean scores of "somewhat worse" and "same" SF2 groups
6MWD Six-minute walk distance
SF36 36 Item Short Form Survey
SGRQ St. George's Respiratory Questionnaire
UCSD SOBQ The University of California, San Diego Shortness of Breath Questionnaire