

RESEARCH

Open Access



# Qualitative and psychometric evaluation of the PROMIS®-Fatigue SF-7a scale to assess fatigue in patients with moderately to severely active inflammatory bowel disease

Brian G. Feagan<sup>1</sup>, William J. Sandborn<sup>2</sup>, Bruce E. Sands<sup>3</sup>, Yan Liu<sup>4</sup>, Marion Vetter<sup>4</sup>, Susan D. Mathias<sup>5</sup>, Kuan-Hsiang Gary Huang<sup>4</sup>, Jewel Johanns<sup>4</sup>, Matthew Germinaro<sup>4</sup> and Chenglong Han<sup>4\*</sup> 

## Abstract

**Background** This study evaluated the content validity and psychometric properties of the Patient-Reported Outcomes Measurement Information System® (PROMIS)-Fatigue Short Form 7a (SF-7a) v1.0 scale to determine its suitability in clinical trials to assess fatigue in patients with moderately to severely active Crohn's disease (CD) and ulcerative colitis (UC).

**Methods** A qualitative interview assessed patients' experience living with CD (N = 20) and UC (N = 19). The contents of the SF-7a scale were cognitively debriefed to evaluate content validity. A psychometric evaluation was performed using data from clinical trials of patients with CD (N = 360) and UC (N = 214). Correlations with Inflammatory Bowel Disease Questionnaire (IBDQ), Crohn's Disease Activity Index (CDAI; CD only), and Mayo score (UC only) determined validity. The Patient Global Impression of Change (PGIC) was used to evaluate reliability and responsiveness to change. Using PGIC as an anchor, a preliminary threshold for clinically meaningful change was identified to define fatigue response in both CD and UC patients.

**Results** All patients reported fatigue as a common symptom. Patients confirmed SF-7a items were relevant to assessing fatigue, instructions and response options were clear, and its 7-day recall period was appropriate. Higher SF-7a scores were associated with higher disease activity (CDAI and Mayo score) and lower health-related quality of life (IBDQ), confirming known groups validity. The correlation of the SF-7a scale was higher with fatigue-related items. ( $r_s \geq -0.70$ ) than with items not directly associated with fatigue. Test-retest reliability was moderate to good (0.54–0.89) among patients with stable disease, and responsiveness to change in disease severity was demonstrated from baseline to Week 12. A  $\geq 7$  point decrease was identified as a reasonable threshold to define clinically meaningful improvement.

These data have been presented in part at The American College of Gastroenterology 2021, Las Vegas, NV, USA. Chenglong Han, MD, PhD.

\*Correspondence:  
Chenglong Han  
CHan3@its.jnj.com

Full list of author information is available at the end of the article



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

**Conclusion** The SF-7a scale is a valid, reliable, and sensitive measure of fatigue in patients with moderately to severely active IBD and can be used to evaluate treatment response.

**Keywords** Patient-reported outcomes Measurement Information System® (PROMIS)-Fatigue Short Form 7a (SF-7a) scale, Crohn's disease, Ulcerative colitis, Fatigue, Disease activity

## Background

Inflammatory bowel disease (IBD) is a group of chronic relapsing disorders that includes Crohn's disease (CD) and ulcerative colitis (UC) and is characterized by a progressive disease course that may include diarrhea, rectal bleeding, abdominal pain, bloody stool, and weight loss. IBD is debilitating and associated with a meaningful reduction in health-related quality of life (HRQoL), work impairment, and fatigue [1].

Fatigue is one of the most prevalent and burdensome symptoms among patients with IBD, with about 80% of patients with active disease reporting significant fatigue [1]. The Selecting Therapeutic Targets in Inflammatory Bowel Disease initiative of the International Organization for the Study of Inflammatory Bowel Diseases acknowledges that fatigue must be factored into the regular assessment of patients with IBD [2]. The pathogenesis of fatigue is complex and remains largely undefined; however, evidence suggests that it may involve physiological, psychological, and social aspects [1, 3]. For instance, chronic fatigue is more prevalent in patients with IBD and has been associated with pain intensity [4] and individual patients with IBD may have different fatigue experiences. This observation underscores the need to evaluate fatigue in patients with IBD. In addition, the content validity of a patient-reported outcome (PRO) instrument must be confirmed to effectively measure fatigue in clinical trials. However, clinical data related to fatigue in IBD patients remain limited.

The Patient-Reported Outcomes Measurement Information System® (PROMIS)-Fatigue Short Form 7a (SF-7a) scale evaluates the experience of fatigue and associated impact on daily life. The objectives of these analyses were to assess the content validity and evaluate the psychometric properties of the SF-7a scale in patients with moderately to severely active CD and UC.

## Methods

### Qualitative interview

Seven clinical sites in the United States (US) identified IBD patients to participate in qualitative interviews to assess the content validity of the SF-7a scale. Eligible patients were required to be adults  $\geq 18$  years of age with a confirmed diagnosis of moderately to severely active CD (defined as Crohn's Disease Activity Index [CDAI]  $\geq 220$  and  $\leq 450$ ) and confirmed by radiography, histology, and/or endoscopy, or patients with confirmed diagnosis of moderately to severely active UC (defined as total

Mayo score of 6 to 12 or a partial Mayo score of  $\geq 5$ ) for at least 3 months. Patients with a medical or psychiatric condition resulting in cognitive or any other impairment were excluded from the study.

Eligible patients were able to speak, read, and write in English and provided written informed consent prior to participating in the interview in which they were asked to complete a background questionnaire, which captured demographic and clinical information, as well as the Inflammatory Bowel Disease Questionnaire (IBDQ). A 1:1 combined-concept elicitation-cognitive debrief semi-structured interview was conducted via Zoom video conference (~60 min in duration). Patients were asked open-ended questions about living with IBD, symptoms they experience, and impacts on their lives using semi-structured interview guides (one for CD patients and one for UC patients) that were designed specifically for this study. Next, a cognitive debriefing interview was conducted in which patients were asked to provide feedback on their interpretation of (and any problems with) the instructions, each item of the SF-7a scale, response options, and recall period using questions, such as "In your own words, what is this question asking you to think about?; How did you respond to this question?; Why did you give that response?; Is this question clear or unclear? Why?"

This approach enabled an evaluation of the relevance, interpretability, clarity, and ease of understanding of all concepts and content of the SF-7a scale. All interviews were recorded and transcribed. Data from the interviews were coded using a qualitative data analysis software, MAXQDA (version 2020). Two coding dictionaries were developed (one for CD patients and one for UC patients) and used in the analysis of the transcripts. The codebook was used to organize and categorize concepts of interest from the interviews and included descriptions and examples for each code to ensure consistency across coders. Each transcript was coded by one coder, reviewed, summarized, and analyzed by a second coder who performed the analysis.

### Psychometric evaluation

#### Patient population

Data for the psychometric evaluation were available from participants enrolled in the GALAXI-1 (NCT03466411) and VEGA (NCT03662542) studies. GALAXI-1, a Phase 2, dose-ranging, double-blind, active- and placebo-controlled study, was designed to evaluate the efficacy

and safety of guselkumab in patients with moderately to severely active CD. VEGA, a Phase 2a, randomized, double-blind, active-controlled, parallel-group, multicenter study, was designed to evaluate the efficacy and safety of combination induction therapy with guselkumab and golimumab in patients with moderately to severely active UC. Detailed descriptions of each study, including eligibility criteria, patient and disease characteristics, and key efficacy and safety outcomes have been presented elsewhere [5–11]. Here, we present results from the psychometric analysis using pooled data across treatment groups by each study.

#### **PROMIS-Fatigue SF-7a v1.0 scale**

The SF-7a scale is a 7-item instrument that evaluates fatigue-related symptoms (i.e., tiredness, exhaustion, mental tiredness, and lack of energy) and associated impacts on daily activities (i.e., activity limitations related to work, self-care, and exercise) during the past 7 days with a 5-point Likert scale that ranges from “1=Never” to “5=Always” (Supplemental Table 1). Higher SF-7a scores indicate more fatigue. The raw score of SF-7a is converted into a standardized T-score based on the US national norm with a mean T-score of 50 and standard deviation (SD) of 10 [12]. Mean T-scores >50 indicate the presence of fatigue.

#### **Anchor assessments for validation**

The IBDQ is a 32-item IBD-specific questionnaire that has been adapted and validated into several languages and cultural milieus [13]. IBDQ is composed of 4 domains (bowel symptoms, systemic symptoms, emotional function, and social function) that are assessed within the past 14 days [14]. IBDQ items are scored on a 7-point Likert scale that ranges from “1=Worst” to “7=Best” yielding a total score that ranges from 32 to 224, with higher IBDQ total scores indicating better HRQoL. The Patient Global Impression of Change (PGIC) is a single item used to evaluate overall change in health status from the patient perspective using a 7-point Likert scale (“A lot better”, “Moderately better”, “A little better”, “Neither better, nor worse (no change)”, “A little worse”, “Moderately worse”, “A lot worse” [15]).

The CDAI score, which is most commonly utilized to assess disease activity in patients with CD, ranges from 0 to 600, with higher CDAI scores indicating more severe active disease [16].

The Mayo score evaluates disease activity in UC patients and is comprised of 4 components: stool frequency, rectal bleeding, endoscopy, and physician’s global assessment [14]. Each is scored on a scale from 0 to 3, and the maximum total score is 12, with higher Mayo scores indicating more severe disease. The partial Mayo score uses 3 non-invasive components of the full Mayo

score (i.e., stool frequency, rectal bleeding, and physician’s global assessment), with a maximum score of 9.

#### **Statistical analysis**

Baseline demographic and clinical characteristics were summarized using descriptive statistics (mean±SD), unless otherwise stated. Known groups validity assesses whether a measure can distinguish groups that are clinically different. Patients were stratified into 3 subgroups based on their HRQOL by IBDQ score [remission (no or mild impact): ≥170; moderate impact:169–110; severe impact: <110] [17, 18], CD disease severity by CDAI score (remission: <150; mild: 150–219; moderate to severe: 220–450; very severe: >450) [19] or UC severity by partial Mayo score (remission: 0-1; mild: 2-4; moderate: 5-6; severe: >6). Mean SF-7a scores were summarized by defined subgroups and compared using analysis of variance.

Convergent validity evaluates whether measures that theoretically should be correlated are, in fact, observed to be highly correlated. Discriminant validity assesses whether measures that theoretically should not be correlated are, in fact, observed to be not correlated. The convergent and discriminant validity were evaluated by comparing Spearman’s correlation between the SF-7a scale and individual IBDQ items related to fatigue/energy level and items less likely related with fatigue.

Test-retest reliability is the process of reproducing identical measurement among stable patients. The intra-class correlation coefficients (ICC) of SF-7a were calculated using data from patients with stable disease defined by PGIC as “no change” from the baseline at Week 12. In general, ICC values below 0.50 indicate poor reliability, 0.50 to 0.75 moderate reliability, 0.75 to 0.90 good reliability, and >0.90 excellent reliability [20].

Responsiveness, which measures how effective a scale detects change over time, was evaluated for the SF-7a scale by calculating the standardized effect size (SES), standardized response mean (SRM), and Guyatt’s statistic. The SES [21] was calculated as the difference in means between baseline and Week 12 in the SF-7a scale divided by the baseline SD. The SRM [22] was calculated as the difference in means between baseline and Week 12 divided by the SD of the change scores. The responsiveness statistic [23] was calculated as the difference in means between baseline and Week 12 in the SF-7a scale divided by the SD of the change in score for stable patients as assessed by the PGIC.

Using PGIC as an anchor variable, changes in the SF-7a scale were summarized across 7 response levels of PGIC. The clinically meaningful threshold was defined as the mean change in SF-7a scores among patients with PGIC response of “moderately better” (representing two levels of improvement) at Week 12 [24].

**Table 1** Patient demographics and disease characteristics for CD and UC patients participating in the qualitative interview

	CD patients (N=20)	UC patients (N=19)
Age, years	43.1 (15.1)	46.1 (12.9)
Male, n (%)	12 (60)	6 (32)
Duration of disease, years	14.9 (9.6)	11.1 (9.0)

Data are being reported as mean (SD), unless otherwise specified. CD, Crohn's disease; UC, ulcerative colitis

### Institutional review board

The study protocol for the qualitative research and informed consent forms were reviewed and approved by WCG Institutional Review Board prior to the start of the study on August 6, 2020.

## Results

### Qualitative interview

A total of 39 patients (20 CD and 19 UC) participated in the qualitative interview. Age and duration of disease were similar between CD and UC patients (Table 1). The majority of CD patients were male (60%) and the mean age was 43 years. Overall, the CD patient population had a mean disease duration of 14.9 years. The UC patient population was predominately female (68%) and the mean age was 46 years. Overall, UC patients had a mean disease duration of 11.1 years.

### Concept elicitation

All patients (100%) in both the CD and UC groups spontaneously reported fatigue as one of their most common symptoms (see patient quotes below). Other symptoms mentioned by patients included increased stool frequency (CD, 100%; UC, 100%), abdominal pain/cramping (CD, 95%; UC, 100%), diarrhea (CD, 100%; UC, 74%), urgency (CD, 80%; UC, 100%), joint pain (CD, 80%; UC, 11%), bloating (CD, 35%; UC, 37%), blood/mucus in stool (CD, 85%; UC, 100%), inability to pass stool despite urgency (CD, 15%; UC, 68%), rectal pain (CD, 10%; UC, 42%), rectal bleeding (CD, 60%; UC, 53%), gas (CD, 40%; UC, 16%), and nausea (CD, 25%; UC, 21%).

### Patient reflections on fatigue

Patients were given an opportunity to describe their fatigue based on their personal experiences. Patients reported that fatigue symptoms impacted their ability to do daily and physical activities. One CD patient said, "For the most part, I'm pretty exhausted all the time. Of course, I have like the body of an 80-year-old, but yeah I mean I'm just wiped out all the time. I probably have to take a nap at least three or four days out of the week because I can't make it through the full day without pain." Another CD patient noted, "Yeah. I'll wake up in the morning and

**Table 2** Cognitive debriefing of the PROMIS Fatigue SF-7a scale

	SF-7a scale	
	CD patients (N=20)	UC patients (N=19)
Accurately paraphrase?	90-100% <sup>a</sup>	90-100% <sup>a</sup>
Questions clear?	78-100% <sup>b</sup>	88-100% <sup>b</sup>
Easy to think about past 7 days?	95% <sup>c</sup>	100%
Questions relevant?	95%	79%
Questionnaire easy to complete?	95%	100%

CD, Crohn's disease; PROMIS, Patient-Reported Outcomes Measurement Information System; SF-7a, Short Form-7a; UC, ulcerative colitis. <sup>a</sup>The number of patients who were asked whether SF-7a questions were accurately paraphrased varied in CD patients from n=9 to n=15 and in UC patients from n=7 to n=15 for each SF-7a question. <sup>b</sup>The number of patients who were asked whether SF-7a questions were clear varied in CD patients from n=18 to n=20 and in UC patients from n=16 to n=18 for each SF-7a question. <sup>c</sup>Based on n=19 patients

*I'll be fine for a couple of hours, and all of a sudden out of nowhere you just feel drained and want to lay back down."*

Similar quotes were provided from UC patients. One UC patient replied, "Oh, lord, yes. You feel very, like you're just destroyed. You know like you are just, there's nothing left. You can't, you're just, like I want to lay down and just rest and you think if I can just rest then I'll feel better and you don't." Another UC patient mentioned, "Slight fatigue on a daily basis, but if it's inflamed most definitely. It's very exhausting. It takes all the life out of me."

When describing fatigue experience, patients with CD or UC often talked about their fatigue intensity/severity and frequency together using expressions such as "constantly exhausted," or "wiped out all the time"; indicating both fatigue frequency and intensity are meaningful attributes to evaluate fatigue experience.

### Cognitive debriefing

Cognitive debriefing of the PROMIS-Fatigue SF-7a revealed that the majority of CD and UC patients (>75%) were able to accurately paraphrase all items, found that items were clear and easy to recall, relevant to fatigue, and easy to complete. Fatigue experience was easy to recall within the previous 7 days (Table 2).

The cognitive debriefing allowed additional information to be gathered from the CD and UC patients based on the patient responses to the SF-7a. For example, 35% of CD patients indicated that work caused them to run out of energy, 64% noted that fatigue limited their job-related responsibilities, 31% reported being too tired to think clearly in the evening/at the end of the day, and 50% described strenuous exercise as lifting weights or doing crunches.

Similarly, 36% of UC patients indicated that daily activities/everything caused them to run out of energy, 60% noted that fatigue limited their work for pay responsibilities, 38% reported being too tired to think clearly at the

end of the day, and 46% described strenuous exercise as weightlifting.

### Psychometric evaluation

Baseline demographics and clinical characteristics of patients with CD and UC are summarized in Table 3. Mean (SD) T-score values of PROMIS-Fatigue SF-7a at baseline were 59.0 (8.2) and 57.4 (8.4) for CD and UC patients, respectively, which were >50 (the population norm) indicating fatigue.

### Known groups validity

CD patients in IBDQ remission (defined as IBDQ total score  $\geq 170$ ) had lower SF-7a values (mean 46.6) than patients with a severe (defined as IBDQ total score <110) IBDQ score (mean 64.1, nominal  $P < 0.001$ ; Fig. 1a). Similarly, UC patients in IBDQ remission had lower SF-7a values (mean 44.8) than patients with a severe IBDQ score (mean 64.3, nominal  $P < 0.001$ ; Fig. 1b). Less fatigue (i.e., lower SF-7a mean scores) was associated with better general health and well-being (i.e., higher IBDQ mean scores). Of note, CD and UC patients in IBDQ remission had SF-7a fatigue mean T-scores <50, indicating that these patients were less impacted by fatigue.

CD patients in CDAI remission (defined as CDAI score <150) had lower SF-7a values (mean 48.6) than patients with a very severe (defined as CDAI score >450) CDAI score (mean 65.7, nominal  $P < 0.001$ ; Fig. 2). Less fatigue (i.e., lower SF-7a scores) was associated with lower disease activity (i.e., lower CDAI scores). CDAI remission was associated with SF-7a fatigue mean T-scores <50, indicating that patients in clinical remission were less impacted by fatigue.

UC patients in remission (defined as partial Mayo score 0–1) had lower SF-7a values (mean 45.8) than patients with severe disease severity (mean 61.9, nominal  $P < 0.001$ ; Fig. 3). UC patients who had moderately

to severely active UC (defined as partial Mayo scores  $\geq 5$ ) had fatigue mean T-scores >50, indicating that these patients were impacted by fatigue.

### Convergent and discriminant validity

Spearman correlations revealed that SF-7a mean scores were strongly correlated ( $r_s \geq -0.70$ ) with fatigue-related IBDQ items (e.g., “feeling of fatigue or tiredness”) in patients with CD and UC (Table 4). IBDQ item “energy levels” was also found to be correlated with SF-7a in CD.

( $r_s = -0.77$ ) and UC ( $r_s = -0.60$ ) patients.

In contrast, relatively weaker correlations were observed with IBDQ items that were not directly fatigue-related, e.g., “accidental soiling of underpants” (CD:  $r_s = -0.44$ ; UC:  $r_s = -0.50$ ) and “rectal bleeding” (CD:  $r_s = -0.26$ ; UC:  $r_s = -0.40$ ). These findings further implied the convergent and divergent validity of the SF-7a scale.

### Test-retest reliability

Test-retest reliability of the SF-7a scale was assessed among stable (defined as no change in PGIC from baseline to Week 12) CD and UC patients (Table 5). Mean SF-7a scores were similar at baseline and Week 12, with ICC values 0.54 in CD and 0.89 in UC. Overall, these findings indicate that the test-retest reliability of the SF-7a scale is adequate.

### Responsiveness and threshold for clinically meaningful change

The SF-7a scale was responsive to change in disease severity from baseline to Week 12 as assessed by the PGIC in both CD and UC patients (Table 6). Using the PGIC as an anchor variable, CD and UC patients with greater improvements in PGIC had greater reductions in SF-7a scores. A 3-level change (i.e., improvement) of feeling “a lot better” was associated with an 11.5-point mean reduction in the SF-7a scale in CD patients and an 11.1-point mean reduction in UC patients. Similarly, a 2-level change (i.e., improvement) of feeling “moderately better” was associated with a 6-point mean reduction in the SF-7a scale in CD patients and a 6.4-point mean reduction in UC patients.

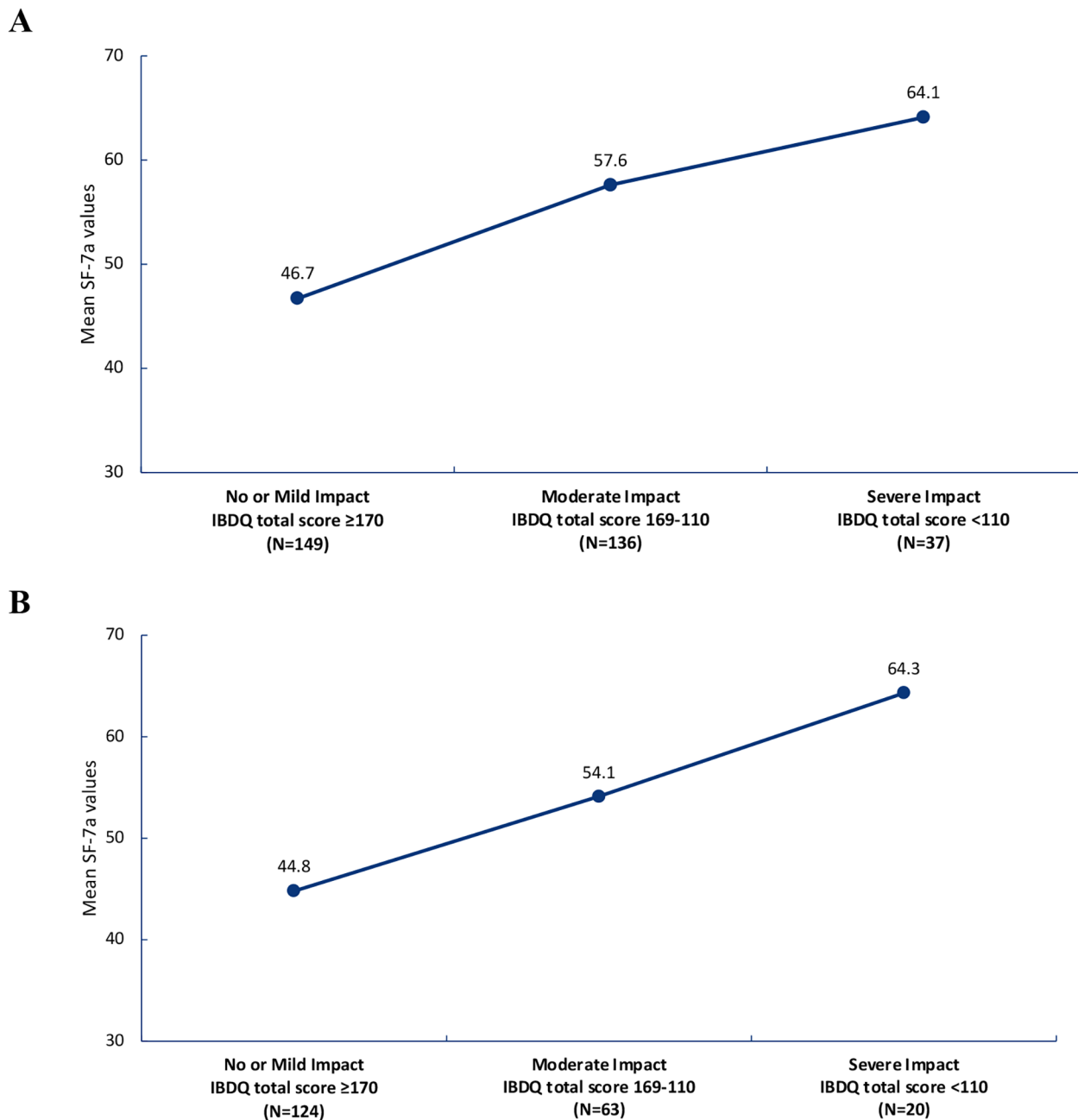
### Discussion

This study evaluated the content validity and psychometric properties of the PROMIS-Fatigue SF-7a v1.0 scale to determine its suitability in clinical trials to assess fatigue in patients with moderately to severely active CD and UC. Our analysis demonstrated that all patients with moderately to severely active CD and UC reported fatigue as a common symptom. Patients who participated in the qualitative interview confirmed that SF-7a items were relevant to assessing fatigue, the instructions and response options were clear, and its 7-day recall period

**Table 3** Patient and disease demographics at baseline in GALAXI (CD patients) and VEGA (UC patients)

	CD patients (N = 360)	UC patients (N = 214)
Age, years	39.3 (13.9)	38.4 (12.0)
Male, n (%)	217 (60.3)	116 (54.2)
Duration of disease, years	8.9 (8.8)	4.9 (5.0)
CDAI total score	305.2 (56.7)	
Mayo score		8.8 (1.4)
IBDQ total score	125.8 (33.3) <sup>a</sup>	116.5 (33.3) <sup>b</sup>
PROMIS Fatigue SF-7a score	59.0 (8.2) <sup>a</sup>	57.4 (8.4) <sup>b</sup>

CD and UC patients were derived from pooled treatment groups from GALAXI and VEGA studies, respectively. Data are being reported as mean (SD), unless otherwise specified. CD, Crohn’s disease; CDAI, Crohn’s Disease Activity Index; IBDQ, Inflammatory Bowel Disease Questionnaire; PROMIS, Patient-Reported Outcomes Measurement Information System; SD, standard deviation; SF-7a, Short Form-7a; UC, ulcerative colitis. <sup>a</sup>Based on 353 CD patients. <sup>b</sup>Based on 209 UC patients

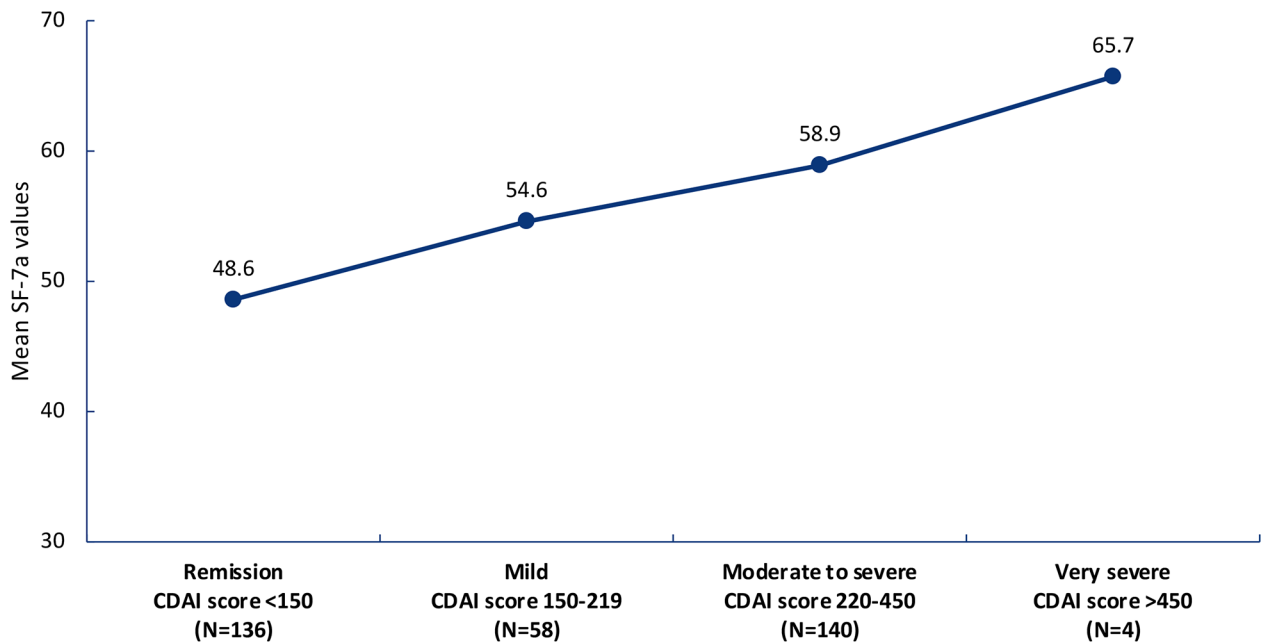


**Fig. 1** Mean PROMIS Fatigue SF-7a values, by impact on HRQoL by IBDQ total score at Week 12 in patients with **(A)** CD and **(B)** UC. CD, Crohn's disease; IBDQ, Inflammatory Bowel Disease Questionnaire; PROMIS, Patient-Reported Outcomes Measurement Information System; SF-7a, Short Form-7a; UC, ulcerative colitis

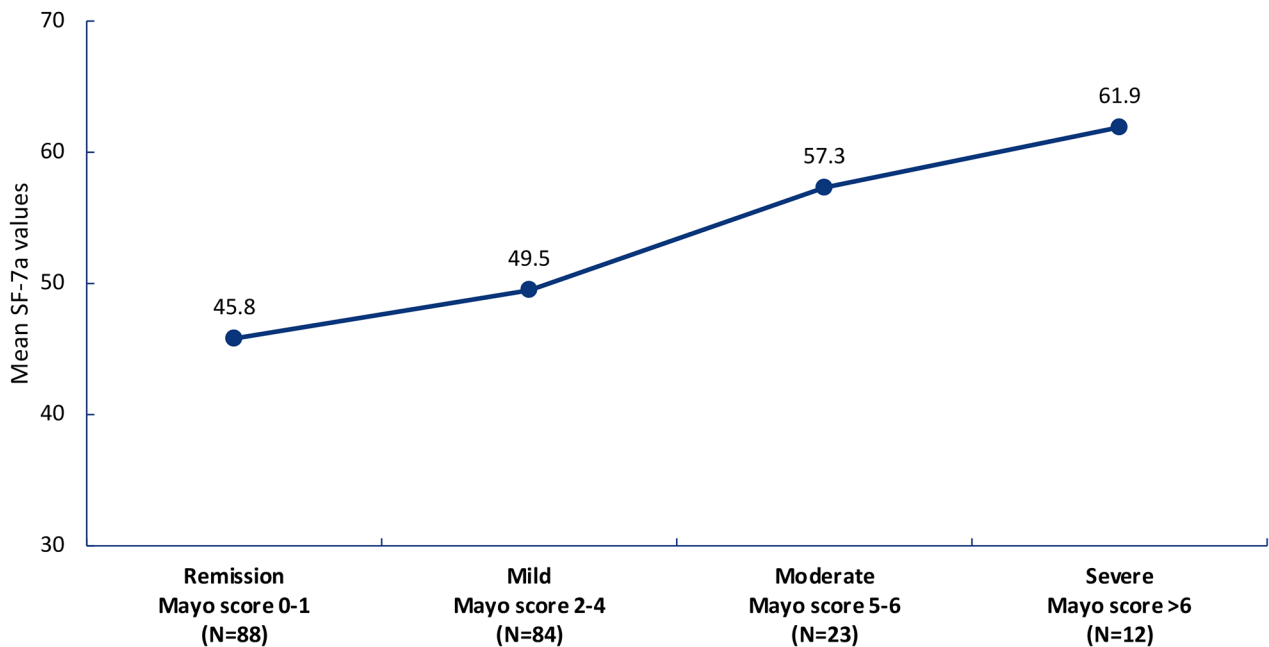
was appropriate. In addition, higher SF-7a scores were associated with higher disease activity (CDAI and Mayo score) and lower HRQoL (IBDQ), confirming known groups validity. The correlation of the SF-7a scale was higher with fatigue-related items ( $r_s \geq -0.70$ ) than with items not directly associated with fatigue. Test-retest reliability was moderate to good (0.54 to 0.89) among patients with stable disease, and responsiveness to

change in disease severity was demonstrated from baseline to Week 12 (e.g., “moderately better” means ~6-point improvement). A  $\geq 7$ -point decrease was identified as a reasonable threshold to define clinically meaningful improvement.

Fatigue in chronic diseases has been clinically described as a “persistent, overwhelming sense of tiredness, weakness, or exhaustion resulting in a decreased



**Fig. 2** Mean PROMIS Fatigue SF-7a values, by disease severity groups by CDAI at Week 12 in patients with CD. CD, Crohn's disease; CDAI, Crohn's Disease Activity Index; PROMIS, Patient-Reported Outcomes Measurement Information System; SF-7a, Short Form-7a



**Fig. 3** Mean PROMIS Fatigue SF-7a values, by disease severity groups by partial Mayo score at Week 12 in patients with UC. PROMIS, Patient-Reported Outcomes Measurement Information System; SF-7a, Short Form-7a; UC, ulcerative colitis

capacity for physical and/or mental work,” and is typically unrelieved by adequate sleep or rest [25]. Furthermore, fatigue can exist as a unique entity and not merely a component of psychological comorbidity or “illness behavior” [25]. The prevalence of fatigue in moderately to severely active CD and UC patients supports the rationale to have

an IBD-specific PRO instrument to assess fatigue in clinical trials.

There are numerous PRO instruments that have been used to measure fatigue in IBD clinical studies. Common fatigue instruments, such as the Multidimensional Fatigue Symptom Inventory-Short Form (MFSI-SF), Brief

**Table 4** Convergent and discriminant analyses: Correlation between PROMIS Fatigue SF-7a scores and IBDQ items at Week 12

IBDQ questions/items	CD	UC
	patients (N=305)	patients (N=204)
Accidental soiling of underpants	-0.44	-0.50
Angry because of bowel problem	-0.54	-0.62
Avoid events where no washroom	-0.57	-0.61
Delay or cancel social engagement	-0.57	-0.60
Difficulty leisure/sports activity	-0.68	-0.62
Embarrassed by bowel problem	-0.59	-0.63
<b>Energy</b>	<b>-0.77</b>	-0.60
Fear of not finding a washroom	-0.54	-0.54
<b>Feeling of fatigue or tiredness</b>	<b>-0.80</b>	<b>-0.70</b>
Felt depressed or discouraged	-0.61	-0.68
Felt generally unwell	-0.65	-0.66
Felt irritable	-0.59	-0.63
Felt relaxed and free of tension	-0.61	-0.59
Felt tearful or upset	-0.52	-0.62
Frustrated, impatient, restless	-0.66	-0.66
Going to bathroom, empty bowels	-0.52	-0.54
Lack of understanding from others	-0.49	-0.62
Limited sexual activity	-0.50	-0.56
Loose bowel movements	-0.44	-0.50
Nausea, feeling sick to stomach	-0.48	-0.57
Problem large amounts of gas	-0.42	-0.53
Problem maintaining/losing weight	-0.51	-0.46
Problems with sleep	-0.62	-0.52
Rectal bleeding	-0.26	-0.40
Satisfied, happy with personal life	-0.53	-0.56
Troubled by abdominal bloating	-0.51	-0.52
Troubled by cramps in abdomen	-0.55	-0.56
Troubled by pain in the abdomen	-0.53	-0.50
Unable to attend school/do work	-0.57	-0.54
Worried or anxious	-0.55	-0.56
Worried, possibility of surgery	-0.45	-0.41

CD, Crohn's disease; IBDQ, Inflammatory Bowel Disease Questionnaire; PROMIS, Patient-reported Outcomes Measurement Information System; SF-7a, Short Form-7a; UC, ulcerative colitis. Strong correlations ( $r_s \geq -0.70$ ) between fatigue-related IBDQ items and mean SF-7a scores are bolded

Fatigue Inventory (BFI), and Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-F), have all been reported to be valid and reliable in assessing fatigue in other disease conditions, but their application remains limited in patients with IBD [26].

MFSI-SF, a 30-item instrument, measures fatigue in 5 multidimensional subscales over the previous week, and appears to be sensitive in detecting change in cancer-related fatigue [27]. BFI assesses fatigue severity and interference with daily function over the past 24 h [28]; however, its test-retest reliability after repeated administrations remains a concern [29]. FACIT-F, a 13-item instrument originally designed to assess cancer-related anemia [30], has been validated in the general population as well as in patients with a number of other disease conditions. However, some items included in the FACIT-F may be interpreted as not relevant to fatigue [31], raising concerns on content validity. During a recent cognitive debriefing of FACIT-Fatigue, CD (N=30) and UC (N=33) patients were asked to report whether anything relevant to their fatigue experience was missing from the FACIT-Fatigue questionnaire [32]. Mental health/mental fatigue that can lead to changes in one's thinking process or mood was reported as a missing concept by 6 patients with CD and 1 patient with UC. Mental fatigue is assessed by one item in PROMIS-Fatigue SF-7a ("How often were you too tired to think clearly?").

This study assessed the content validity and psychometric properties of the SF-7a scale in patients with IBD. In the qualitative analysis, all CD and UC patients identified fatigue as the most common symptom, with all patients having spontaneously reported experiencing fatigue. Patients with CD or UC often talked about their fatigue intensity/severity and fatigue frequency together using expressions such as "constantly exhausted," or "wiped out all the time", indicating both fatigue frequency and intensity are important attributes in the evaluation of fatigue experience.

Cognitive debriefing of the SF-7a scale confirmed the content validity of the SF-7a scale and showed that most patients correctly interpreted each item, felt each item was relevant to their experience, and found the questionnaire was easy to recall and complete. The SF-7a scale was designed to assess fatigue severity from general

**Table 5** Reliability of PROMIS Fatigue SF-7a scores among stable patients at baseline and Week 12

PROMIS Scale	No change in PGIC from baseline to Week 12 in patients with CD (N=74)			No change in PGIC from baseline to Week 12 in patients with UC (N=15)		
	Baseline Mean (SD)	Week 12 Mean (SD)	ICC (95% CI)	Baseline Mean (SD)	Week 12 Mean (SD)	ICC (95% CI)
Fatigue SF-7a	59.6 (7.0)	58.0 (7.9)	0.54 (0.38, 0.69)	62.4 (10.4)	63.2 (10.1)	0.89 (0.74, 0.96)

CD, Crohn's disease; CI, confidence interval; ICC, intraclass correlation coefficient; PGIC, Patient Global Impression of Change; PROMIS, Patient-reported Outcomes Measurement Information System; SD, standard deviation; SF-7a, Short Form-7a; UC, ulcerative colitis



**Table 6** Responsiveness of PROMIS Fatigue SF-7a scale: Change from baseline in fatigue score at Week 12

PGIC from baseline to Week 12	CD					UC				
	N	Mean (SD)	SES	SRM	Guyatt's	N	Mean (SD)	SES	SRM	Guyatt's
A lot better	71	-11.5 (8.5)	-1.6	-1.4	-1.6	106	-11.1 (8.4)	-1.3	-1.3	-1.1
Moderately better	63	-6.0 (9.1)	-0.7	-0.7	-0.9	43	-6.4 (8.0)	-0.9	-0.8	-0.6
A little better	94	-4.3 (6.9)	-0.5	-0.6	-0.6	34	-5.7 (8.1)	-0.6	-0.7	-0.6
No change	74	-1.6 (7.0)	-0.2	-0.2	-0.2	15	0.9 (4.7)	0.1	0.2	0.1
A little worse to a lot worse	15	-2.4 (11.1)	-0.3	-0.2	-0.3	6	-0.1 (7.9)	-0.03	-0.01	-0.01

CD, Crohn's disease; Guyatt's, Guyatt's responsiveness statistic; PGIC, Patient Global Impression of Change; PROMIS, Patient-Reported Outcomes Measurement Information System; SD, standard deviation; SES, standardized effect size; SF-7a, Short Form-7a; SRM, standardized response mean; UC, ulcerative colitis

tiredness and lack of energy to exhaustion and mental tiredness with a response option on a 5-point Likert scale to capture frequency over the past 7 days. Assessing both fatigue intensity and frequency using only 7 items may make the SF-7a scale more valuable in evaluating fatigue in the IBD population.

We have also demonstrated that the SF-7a scale was reliable, valid, and responsive to change in patients with moderately to severely active CD and UC. The SF-7a scale was sensitive in distinguishing clinically different groups as defined by IBDQ, CDAI, and partial Mayo score. The SF-7a scale had adequate reliabilities, and as expected, the strongest correlations ( $r_s \geq -0.70$ ) were identified with fatigue-related IBDQ items versus relatively weaker correlations with other IBDQ items not directly assessing fatigue. Compared to the weak correlation of FACIT-F with clinical outcomes in CD patients compared with UC patients as described in a previous study [26], the current psychometric analysis revealed that SF-7a demonstrated similar correlations with clinical outcomes, regardless of UC or CD disease indication. These findings contribute to the supportive clinical evidence for the use of the PROMIS-Fatigue SF-7a scale in clinical trials for patients with CD and UC.

While the SF-7a scale has been accepted by the Food and Drug Administration to evaluate the treatment effect on fatigue in patients with myelofibrosis [33], the qualitative and quantitative analysis presented herein may provide evidence supporting SF-7a scale as a valid instrument to evaluate fatigue for disease state and progression for patients with CD and UC. One strength of our analysis was that the patients enrolled in both qualitative and quantitative studies had similar disease severity at baseline (moderately to severely active). Also, anchor-based analyses were used to assess changes in clinical outcomes (e.g., HRQoL, disease activity, and disease severity) as they relate to mean SF-7a fatigue values, both of which meet the general requirements for developing a PRO instrument for inclusion in clinical trials that are fit-for-purpose and meaningful to patients, clinicians, regulators, and payers [34]. Moreover, it was identified that a 6- to 6.4-point reduction in the SF-7a scale was clinically

meaningful, which may help facilitate informed decisions in establishing appropriate response criteria for detecting clinically meaningful thresholds with the SF-7a scale. Since two levels of improvement in PGIC was generally considered a clinically meaningful change in disease severity from a patient perspective, the observed change in SF-7a among those patients was defined as meaningful change thresholds. Therefore, a  $\geq 7$ -point change is a preliminary estimate and can be used as a conservative cutoff to define fatigue response in both CD and UC patients.

A potential limitation of this study is that the Phase 2 clinical trials which provided the data for evaluation of the measurement properties enrolled a relatively homogenous patient population, since all patients were required to have moderately to severely active disease. Fatigue assessment by SF-7a may be limited among patients with moderately to severely active disease, additional evaluation of fatigue using SF-7a in patients with mild disease is needed. In addition, due to the limited sample size, there were no suitable analyses to determine whether the SF-7a scale was responsive to worsening of disease. Other limitations include the patient population in the qualitative study being restricted to US patients, the time period in which the reliability assessments were conducted, and the lack of specific hypotheses for correlations between overall disease severity and systemic fatigue. Shorter time intervals between test-retest may prove to yield improved reliabilities, particularly in CD patients. Further validation studies of IBD patients from real-world settings are needed and responsiveness to efficacious treatment vs. placebo needs to be further tested.

In conclusion, we have demonstrated that the SF-7a scale is a valid, reliable, and sensitive measure of fatigue in patients with moderately to severely active CD or UC and can be used to evaluate treatment response. These findings provide supportive evidence for the use of the SF-7a scale in clinical studies of patients with moderately to severely active CD and UC. These data may be helpful in informing future clinical trials as clinicians seek to understand how potential treatment options in IBD may impact fatigue.

## Abbreviations

BFI	Brief Fatigue Inventory
CD	Crohn's disease
CDAI	Crohn's Disease Activity Index
FACIT-F	Functional Assessment of Chronic Illness Therapy-Fatigue
HRQoL	health-related quality of life
IBD	inflammatory bowel disease
IBDQ	Inflammatory Bowel Disease Questionnaire
ICC	intraclass correlation coefficients
MFSI-SF	Multidimensional Fatigue Symptom Inventory-Short Form
PGIC	Patient Global Impression of Change
PRO	Patient-Reported Outcome
PROMIS	Patient-Reported Outcomes Measurement Information System®
SD	standard deviation
SES	standardized effect size
SF-7a	Short Form-7a
SRM	standardized response mean
UC	ulcerative colitis
US	United States

## Acknowledgements

Medical writing support was provided by Chastity Bradley, PhD, of Synchrogenix, a Certara Company, under the direction of the authors in accordance with Good Publication Practice guidelines and was funded by Janssen Research & Development, LLC, and Janssen Scientific Affairs, LLC.

## Author contribution

CH was involved in study concept and design. All authors were involved in the acquisition of data. YL, SDM, and CH were involved in the analysis of data. All authors were involved in the interpretation of data and in drafting of the manuscript. YL and CH were responsible for providing administrative, technical, or material support. CH provided study supervision. All authors approved the final version of the manuscript. YL, MV, K-HGH, JJ, MG, and CH assisted in creating the PROMIS-Fatigue SF-7a instrument.

## Funding

This study was sponsored by Janssen Research & Development, LLC and Janssen Scientific Affairs, LLC.

## Data Availability

The request to share data underlying this article will be individually assessed upon reasonable request to the corresponding author.

## Declarations

### Ethics approval and consent to participate

The study protocol for the qualitative research and informed consent forms were reviewed and approved by an independent scientific review committee prior to the start of the study. Similarly, the study protocols for the GALAXI and VEGA studies were approved by the Institutional Review Board or Independent Ethics Committee at each participating investigative center. All patients provided written informed consent prior to participating in the qualitative interview and prior to enrollment in the GALAXI and VEGA trials.

### Consent for publication

All authors have reviewed the final manuscript and provided their consent for publication.

### Competing interests

BGF received grant/research support from AbbVie Inc, Amgen Inc, AstraZeneca/MedImmune Ltd, Atlantic Pharmaceuticals Ltd, Boehringer Ingelheim, Celgene Corporation, Celltech, Genentech Inc/Hoffmann-La Roche Ltd, Gilead Sciences Inc, GlaxoSmithKline (GSK), Janssen Research & Development LLC, Pfizer Inc, Receptos Inc/Celgene International, Sanofi, Santarus Inc, Takeda Development Center Americas Inc, Tillotts Pharma AG, and UCB; was a consultant for Abbott/AbbVie, AdMirX Inc, Akebia Therapeutics, Allergan, Amgen, Applied Molecular Transport Inc, Aptevo Therapeutics, Asta Pharma, AstraZeneca, Atlantic Pharma, Avir Pharma, Biogen Idec, BioMx Israel, Boehringer Ingelheim, Boston Pharmaceuticals, Bristol-Myers Squibb, Calypso Biotech, Celgene, Elan/Biogen, EnGene, Ferring Pharma, Genentech/Roche, Galapagos, Galen/Atlantica, GiCare Pharma,

Gilead, Gossamer Pharma, GSK, Inception IBD Inc, Intact Therapeutics, Johnson & Johnson/Janssen, Kyowa Kakko Kirin Co Ltd, Lexicon, Lilly, Lycera BioTech, Merck, Mesoblast Pharma, Millennium, Nestle, Nextbiotix, Novonordisk, ParImmune, Parvus Therapeutics Inc, Pfizer Inc, Prometheus Therapeutics and Diagnostics, Progenity, Protagonist, Qu Biologics, Rebiotix, Receptos, Salix Pharma, Shire, Sienna Biologics, Sigmoid Pharma, Sterna Biologicals, Synergy Pharma Inc, Takeda, Teva Pharma, TiGenix, Tillotts, UCB Pharma, Vertex Pharma, Vivelix Pharma, VHSquared Ltd, and Zyngenia; was a member of the speakers bureau for Abbott/AbbVie, Johnson & Johnson/Janssen, Lilly, Takeda, Tillotts, and UCB Pharma; was a member of the scientific advisory board for Abbott/AbbVie, Allergan, Amgen, AstraZeneca, Atlantic Pharma, Avaxia Biologics Inc, Boehringer Ingelheim, Bristol-Myers Squibb, Celgene, Centocor Inc, Elan/Biogen, Galapagos, Genentech/Roche, Johnson & Johnson/Janssen, Merck, Nestle, Novartis, Novonordisk, Pfizer Inc, Prometheus Laboratories, Protagonist, Salix Pharma, Sterna Biologicals, Takeda, Teva, TiGenix, Tillotts Pharma AG, and UCB Pharma; and is a Senior Scientific Officer - Robarts Clinical Trials Inc. WJS reports research grants from Abbvie, Abivax, Arena Pharmaceuticals, Boehringer Ingelheim, Bristol-Meyers Squibb, Genentech/Roche, Gilead Sciences, Inc., GSK, Janssen, Lilly, Pfizer, Prometheus Laboratories, Seres Therapeutics, Shire Pharmaceuticals, Takeda, Theravance Biopharma; consulting fees from Abbvie, Abivax, Alfasigma, Alimentiv (previously Robarts Clinical Trials, owned by Alimentiv Health Trust), Allakos, Amgen, Arena Pharmaceuticals, AstraZeneca, Atlantic Pharmaceuticals, Beigene, Boehringer Ingelheim, Bristol-Meyers Squibb, Celltrion, Clostrabio, Forbion, Galapagos, Genentech/Roche, GSK, Gossamer Bio, Index Pharmaceuticals, Iota Biosciences, Janssen, Lilly, Morphic Therapeutics, Novartis, Oppilan Pharma (now Ventyx Biosciences), Pfizer, Pharm Olam, Polpharm, Progenity, Prometheus Biosciences, Protagonist Therapeutics, PTM Therapeutics, Seres Therapeutics, Shoreline Biosciences, Sublimity Therapeutics, Surrozen, Takeda, Theravance Biopharma, Vedanta Biosciences, Ventyx Biosciences, Vimalan Biosciences, Vivreon Gastrosciences, Xencor, Zealand Pharmaceuticals; stock or stock options from Allakos, BeiGene, Gossamer Bio, Oppilan Pharma (now Ventyx Biosciences), Progenity, Prometheus Biosciences, Prometheus Laboratories, Protagonist Therapeutics, Shoreline Biosciences, Ventyx Biosciences, Vimalan Biosciences, Vivreon Gastrosciences; and employee at Shoreline Biosciences and Ventyx Biosciences. Spouse: Iveric Bio - consultant, stock options; Progenity - stock; Oppilan Pharma (now Ventyx Biosciences) - stock; Prometheus Biosciences - employee, stock, stock options; Prometheus Laboratories - stock, stock options, consultant; Ventyx Biosciences - stock, stock options; Vimalan Biosciences - stock, stock options. BES reports grants, personal fees, and other from Abivax during the conduct of the study; personal fees from Abbvie; personal fees from Alimentiv (previously Robarts Clinical Trials; owned by Alimentiv Health Trust); personal fees from Amgen; personal fees and other from Arena Pharmaceuticals; personal fees from AstraZeneca; personal fees from Bausch Health (Salix Pharmaceuticals); personal fees from Boehringer Ingelheim; personal fees from Boston Pharmaceuticals; personal fees and other from Bristol-Myers Squibb; personal fees and other from Celgene; personal fees from Celltrion; personal fees from Clostrabio; personal fees from Equillium; personal fees from Galapagos; personal fees from Genentech/Roche; personal fees from Gilead Sciences; personal fees from GSK; personal fees from Gossamer Bio; personal fees from Immunic; personal fees from Index Pharmaceuticals; personal fees from Inotrem; grants, personal fees, and other from Janssen; grants, personal fees, and other from Lilly; personal fees from Morphic Therapeutics; personal fees from Novartis; personal fees from Oppilan Pharma (acquired by Ventyx Biosciences); personal fees from Otuska; grants, personal fees, and other from Pfizer; personal fees from Progenity; personal fees from Prometheus Biosciences; personal fees from Prometheus Laboratories; personal fees from Protagonists Therapeutics; personal fees from Seres Therapeutics; personal fees from Shire Pharmaceuticals; personal fees from Surrozen; grants, personal fees, and other from Takeda; grants and personal fees from Theravance Biopharma; personal fees from Teva; personal fees from UCB; personal fees from Vedanta Biosciences; personal fees and other from Ventyx Biosciences; personal fees from Vivelix Pharmaceuticals; personal fees from Ironwood Pharmaceuticals; personal fees from OSE Immunotherapeutics; personal fees from Redhill Biopharma; personal fees from Target RWE; personal fees from Bacainn Therapeutics; personal fees from Capella Bioscience; personal fees from Palatin Technologies; personal fees from Rheos Medicines; personal fees from USWM Enterprises; personal fees from Viela Bio; personal fees from Artugen Therapeutics; personal fees from Calibr; personal fees from Entera; personal fees from Evommune; personal fees from Innovation Pharmaceuticals; personal fees from Kaleido; personal fees from Merck;

personal fees from MRM Health; personal fees from Q32 Bio; personal fees from Sun Pharma; personal fees from TLL Pharmaceutical; and other from Vivante Health outside the submitted work.

SDM is an employee at Health Outcomes Solutions, which received research funding from Janssen Pharmaceuticals for conducting the qualitative research portion of this study.

YL (at the time study was conducted), MV, K-HGH, JJ, MG, and CH are employees of Janssen Research & Development, LLC and own stock/stock options.

#### Author details

<sup>1</sup>Western University and Alimentiv Inc, London, ON, Canada

<sup>2</sup>University of California San Diego, La Jolla, CA, USA

<sup>3</sup>Icahn School of Medicine at Mount Sinai, New York, NY, USA

<sup>4</sup>Janssen Research & Development, LLC, Spring House, PA, USA

<sup>5</sup>Health Outcomes Solutions, Winter Park, FL, USA

Received: 13 February 2023 / Accepted: 9 October 2023

Published online: 14 November 2023

#### References

1. Borren NZ, van der Woude CJ, Ananthkrishnan AN (2019) Fatigue in IBD: epidemiology, pathophysiology and management. *Nat Rev Gastroenterol Hepatol* 16(4):247–259. <https://doi.org/10.1038/s41575-018-0091-9>
2. Turner D, Ricciuto A, Lewis A, D'Amico F, Dhaliwal J, Griffiths AM et al (2021) STRIDE-II: an update on the selecting therapeutic targets in inflammatory bowel Disease (STRIDE) Initiative of the International Organization for the study of IBD (IOIBD): determining therapeutic goals for treat-to-target strategies in IBD. *Gastroenterology* 160(5):1570–1583. <https://doi.org/10.1053/j.gastro.2020.12.031>
3. Kreijne JE, Lie MR, Vogelaar L, van der Woude CJ (2016) Practical Guideline for Fatigue Management in inflammatory bowel Disease. *J Crohns Colitis* 10(1):105–111. <https://doi.org/10.1093/ecco-jcc/jjv168>
4. Jelsness-Jorgensen LP, Bernklev T, Henriksen M, Torp R, Moum BA (2011) Chronic fatigue is more prevalent in patients with inflammatory bowel Disease than in healthy controls. *Inflamm Bowel Dis* 17(7):1564–1572. <https://doi.org/10.1002/ibd.21530>
5. Sandborn WJ, Chan D, Johanns J, Lang G, Adedokun OJ, Afzali A et al (2020) The efficacy and safety of guselkumab induction therapy in patients with moderately to severely active Crohn's Disease: Week 12 interim analyses from the phase 2 GALAXI-1 study [abstract]. Virtual Meeting. United European Gastroenterology Week (UEGW)
6. D'Haens G, Rubin DT, Panés J, Gonzalez S, Chan D, Johanns J et al (2021) The effect of guselkumab induction therapy on endoscopic outcome measures in patients with moderately to severely active Crohn's Disease: Week 12 results from the phase 2 GALAXI-1 study [abstract]. Digestive Disease Week (DDW). Virtual Meeting
7. Sands BE, Danese S, Andrews JM, Weisel K, Chan D, Johanns J et al (2021) The effect of guselkumab induction therapy on inflammatory biomarkers in patients with moderately to severely active Crohn's Disease: Week 12 results from the phase 2 GALAXI-1 study [abstract]. Digestive Disease Week (DDW). Virtual Meeting
8. Feagan BG, Sands B, Sandborn WJ, Shipitofsky N, Marko M, Sheng S et al (2022) Efficacy and safety of combination induction therapy with guselkumab and golimumab in participants with moderately-to-severely active ulcerative colitis: Results through Week 12 of the phase 2a VEGA study. *Digestive Disease Week, San Diego, CA, 21–24 May 2022*
9. Sands BE, Feagan B, Sandborn WJ, Shipitofsky N, Marko M, Sheng S et al (2022) Efficacy and safety of combination induction therapy with guselkumab and golimumab in participants with moderately-to-severely active ulcerative colitis: Results through Week 12 of a phase 2a randomized, double-blind, active-controlled, parallel-group, multicenter, proof-of-concept study. European Crohn's and Colitis Organisation (ECCO). Virtual 16–19 Feb 2022
10. Sandborn WJ, D'Haens GR, Reinisch W, Panés J, Chan D, Gonzalez S et al (2022) Guselkumab for the treatment of Crohn's disease: Induction results from the phase 2 GALAXI-1 study. *Gastroenterology* 162(6):1650–1664e8. <https://doi.org/10.1053/j.gastro.2022.01.047>
11. Feagan BG, Sands B, Sandborn WJ, Germinaro M, Vetter M, Shao J et al (2023) Guselkumab plus golimumab combination therapy versus guselkumab or golimumab monotherapy in patients with ulcerative colitis (VEGA): A randomized, double-blind, controlled, phase 2, proof-of-concept trial. Epub ahead of print. [https://doi.org/10.1016/S2468-1253\(22\)00427-7](https://doi.org/10.1016/S2468-1253(22)00427-7), 1 Feb 2023
12. PROMIS (2022) [Available from: <https://www.healthmeasures.net/score-and-interpret/interpret-scores/promis>. Accessed 14 Feb 2022
13. Pallis AG, Mouzas IA, Vlachonikolis IG (2004) The inflammatory bowel Disease questionnaire: a review of its national validation studies. *Inflamm Bowel Dis* 10(3):261–269. <https://doi.org/10.1097/00054725-200405000-00014>
14. Clinical Review Report : Adalimumab (Humira) (2016). Canadian Agency for Drugs and Technologies in Health, Ottawa (ON)
15. Boomershine CS (2012) A comprehensive evaluation of standardized assessment tools in the diagnosis of fibromyalgia and in the assessment of fibromyalgia severity. *Pain Res Treat* 2012:653714. <https://doi.org/10.1155/2012/653714>
16. Liu JK, Lichenstein GR (2012) Inflammatory bowel Disease. In: Ginsberg GG, Gostout CJ, Kochman ML, Norton ID (eds) *Clinical gastrointestinal endoscopy*, 2nd edn. Elsevier Saunders, St. Louis, pp 243–264
17. Irvine E (1999) Development and subsequent refinement of the inflammatory bowel Disease Questionnaire: a quality-of-life instrument for adult patients with inflammatory bowel Disease. *J Pediatr Gastroenterol Nutr* 28:S23–S27. <https://doi.org/10.1097/00005176-199904001-00003>
18. Irvine EJ, Yeh C-H, Ramsey D, Stirling AL, Higgins PDR (2008) The effect of mesalazine therapy on quality of life in patients with mildly and moderately active ulcerative Colitis. *Aliment Pharmacol Ther* 28:1278–1286
19. Best WR, Becktel JM, Singleton JW, Kern F Jr (1976) Development of a crohn's disease activity index: National cooperative crohn's disease study. *Gastroenterology* 70(3):439–444
20. Koo TK, Li MY (2016) A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med* 15(2):155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
21. Cohen J (1988) *Statistical power analysis for the behavioral sciences*, 2nd edn. Academic Press, New York
22. Stucki G, Liang MH, Fossel AH, Katz JN (1995) Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar spinal stenosis. *J Clin Epidemiol* 48(11):1369–1378. [https://doi.org/10.1016/0895-4356\(95\)00054-2](https://doi.org/10.1016/0895-4356(95)00054-2)
23. Guyatt GH, Bombardier C, Tugwell PX (1986) Measuring disease-specific quality of life in clinical trials. *CMAJ* 134(8):889–895
24. Critical Path Institute (C-PATH). Using patient input to estimate clinically meaningful within-patient change at the scale score level. Presented at: 10th Annual Patient-Reported Outcome Consortium Workshop; 24–25 April 2019, Silver Spring, MD. Available at: [https://c-path.org/wp-content/uploads/2019/05/2019\\_session6\\_meaningfulchgfinal.pdf](https://c-path.org/wp-content/uploads/2019/05/2019_session6_meaningfulchgfinal.pdf). Accessed 17 August 2023
25. van Langenberg DR, Gibson PR (2010) Systematic review: fatigue in inflammatory bowel Disease. *Aliment Pharmacol Ther* 32(2):131–143. <https://doi.org/10.1111/j.1365-2036.2010.04347.x>
26. Tinsley A, Macklin EA, Korzenik JR, Sands BE (2011) Validation of the Functional Assessment of Chronic Illness therapy-fatigue (FACIT-F) in patients with inflammatory bowel Disease. *Aliment Pharmacol Ther* 34:1328–1336
27. Stein KD, Jacobsen PB, Blanchard CM, Thors C (2004) Further validation of the multidimensional fatigue symptom inventory-short form. *J Pain Symptom Manage* 27(1):14–23. <https://doi.org/10.1016/j.jpainsymman.2003.06.003>
28. Mendoza TR, Wang XS, Cleeland CS, Morrissey M, Johnson BA, Wendt JK et al (1999) The rapid assessment of fatigue severity in cancer patients: use of the brief fatigue inventory. *Cancer* 85(5):1186–1196. [https://doi.org/10.1002/\(sici\)1097-0142\(19990301\)85:5%3C1186::aid-cnrc24%3E3.0.co;2-n](https://doi.org/10.1002/(sici)1097-0142(19990301)85:5%3C1186::aid-cnrc24%3E3.0.co;2-n)
29. Pallett E, Rentowl P, Hanning C (2009) The brief fatigue inventory: comparison of data collection using a novel audio device with conventional paper questionnaire. *J Pain Symptom Manage* 38(3):390–400. <https://doi.org/10.1016/j.jpainsymman.2008.11.015>
30. Yellen SB, Cella DF, Webster K, Blendowski C, Kaplan E (1997) Measuring fatigue and other anemia-related symptoms with the Functional Assessment of Cancer Therapy (FACT) measurement system. *J Pain Symptom Manage* 13(2):63–74. [https://doi.org/10.1016/s0885-3924\(96\)00274-6](https://doi.org/10.1016/s0885-3924(96)00274-6)
31. Kosinski M, Gajria K, Fernandes AW, Cella D (2013) Qualitative validation of the FACIT-fatigue scale in systemic Lupus Erythematosus. *Lupus* 22(5):422–430. <https://doi.org/10.1177/0961203313476360>
32. Loftus EV Jr., Ananthkrishnan AN, Lee WJ, Gonzalez YS, Fitzgerald KA, Wallace K et al (2023) Content validity and psychometric evaluation of the functional

assessment of chronic illness therapy-fatigue (FACIT-Fatigue) in patients with crohn's disease and ulcerative colitis. *Pharmacoecon Open*. <https://doi.org/10.1007/s41669-023-00419-w>

33. Jakafi® (ruxolitinib) tablets. Prescribing information (2011) Incyte Corporation, Wilmington, DE
34. Patient-reported outcome measures: Use in medical product development to support labeling claims guidance for industry (2009) US Department of Health and Human Services, Food and Drug Administration, Center for Drug

Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH), Rockville, MD

### **Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.