



OPEN ACCESS

EDITED BY

Shirley D'Sa,
University College London Hospitals NHS
Foundation Trust, United Kingdom

REVIEWED BY

Maxime Desmarests,
University of Franche-Comté, France
Sigbjørn Berentsen,
Fonna Hospital Trust, Norway

*CORRESPONDENCE

David Cella

✉ d-cella@northwestern.edu

RECEIVED 02 September 2024

ACCEPTED 04 February 2025

PUBLISHED 27 February 2025

CITATION

Cella D, Hill QA, Karaoui A, Joly F,
Sourdille T, Shafer F, Wardęcki M, Kosinski M
and Röth A (2025) FACIT-Fatigue scale in
patients with cold agglutinin disease:
psychometric validation and estimation
of clinically meaningful change.
Front. Hematol. 4:1490130.
doi: 10.3389/frhem.2025.1490130

COPYRIGHT

© 2025 Cella, Hill, Karaoui, Joly, Sourdille,
Shafer, Wardęcki, Kosinski and Röth. This is an
open-access article distributed under the terms
of the [Creative Commons Attribution License
\(CC BY\)](#). The use, distribution or reproduction
in other forums is permitted, provided the
original author(s) and the copyright owner(s)
are credited and that the original publication
in this journal is cited, in accordance with
accepted academic practice. No use,
distribution or reproduction is permitted
which does not comply with these terms.

FACIT-Fatigue scale in patients with cold agglutinin disease: psychometric validation and estimation of clinically meaningful change

David Cella ^{1*}, Quentin A. Hill², Alia Karaoui³, Florence Joly⁴,
Timothee Sourdille⁵, Frank Shafer⁶, Marek Wardęcki⁷,
Mark Kosinski⁸ and Alexander Röth⁹

¹Department of Medical Social Sciences, Feinberg School of Medicine, Northwestern University, Chicago, IL, United States, ²Department of Clinical Hematology, Leeds Teaching Hospitals National Health Service (NHS) Trust, Leeds, United Kingdom, ³BMAPS SARM, Geneva, Switzerland, ⁴Global Market Access, Health Economics & Value Assessments, Sanofi, Chilly-Mazarin, France, ⁵Global Medical, Sanofi, Bridgewater, NJ, United States, ⁶Clinical Research, Sanofi Bridgewater, NJ, United States, ⁷Hematology, Sanofi, Warsaw, Poland, ⁸QualityMetric, Johnston, RI, United States, ⁹Department of Hematology and Stem Cell Transplantation, West German Cancer Center, University Hospital Essen, Essen, Germany

Purpose: To validate the use of Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) questionnaire in cold agglutinin disease (CAD) patients using qualitative and quantitative methods and to estimate meaningful within-patient change (MWPC).

Methods: Qualitative assessment used outcomes from a survey among CAD patients and their caregivers in US. Quantitative assessment used outcomes from two Phase-3 trials in CAD wherein fatigue was evaluated as a key secondary endpoint using the FACIT-Fatigue questionnaire. The reliability, validity, and responsiveness of the FACIT-Fatigue questionnaire were assessed. MWPC was estimated using anchor-based (mean change, receiver operating characteristic [ROC] curves, and logistic regression) and distribution-based methods.

Results: Qualitative analyses ($n=16$) showed that fatigue was the most common and bothersome symptom. All patients reported that FACIT-Fatigue questionnaire captured their experiences of CAD-related fatigue. Quantitative analysis included 55 patients from both studies. Items of FACIT-Fatigue scale were internally consistent (Cronbach's alpha coefficient: 0.94 at baseline; 0.96 at Week 26). Generally, correlations showed good convergent validity (>0.40). The MWPC estimates ranged from 2.0 to 15.7. Based on more robust ROC and regression-based methods, IQR of MWPC estimates was 4.1–7.3, and individual responder definitions were in range of 5–8 points, where "5" is the lowest recommended MWPC threshold for FACIT-Fatigue in CAD.

Conclusion: FACIT-Fatigue is a reliable, valid, responsive scale in CAD. The MWPC estimates for FACIT-Fatigue in patients with CAD were consistent with other disease estimates published previously, and “5” can be considered as the lowest recommended threshold for meaningful clinical response in patients with CAD.

KEYWORDS

Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue), cold agglutinin disease (CAD), Meaningful Within Patient Change (MWPC), FACIT-Fatigue questionnaire, psychometric validation

Introduction

Cold agglutinin disease (CAD) is an autoimmune disease that involves immunoglobulin M (IgM) antibodies and red blood cells (RBCs). When these autoantibodies are exposed to temperatures lower than normal body temperature ($\leq 37^{\circ}\text{C}$), they bind to the I antigen on RBCs, resulting in agglutination and activation of the classical complement pathway (a part of the innate immune system enabling protection against pathogens). Complement activation leads to the premature destruction of RBCs (hemolysis) and then to anemia (1–3). Hemoglobin (Hb) levels are a key clinical marker of CAD severity (4). A nationwide cohort study from Denmark reported a median survival time of 8.8 years in patients with CAD (5).

Patients with CAD experience a variety of complement-mediated chronic hemolysis-related symptoms, such as fatigue, dizziness, headache, shortness of breath, muscle weakness, and jaundice, as well as cold-induced circulatory symptoms, such as acrocyanosis and Raynaud’s phenomenon (6). Profound and chronic fatigue is one of the most common symptoms of CAD and severely affects patients’ quality of life (6, 7). In an internet-based survey among patients with CAD in the United States (US), most patients reported having experienced fatigue (90% [45/50]) with related symptoms of weakness, difficulty concentrating, and headaches (8). Patients with CAD who experienced fatigue found it difficult to complete daily tasks or start new projects, as well as doing household chores (8).

The Functional Assessment of Chronic Illness Therapy-Fatigue (FACIT-Fatigue) is a generic scale used to assess fatigue in chronic diseases by exploring aspects of a patient’s daily life related to personal, professional, social, physical, and emotional well-being. This study aimed to evaluate the psychometric properties and estimate clinically meaningful change for FACIT-Fatigue in CAD.

Abbreviations: CAD, Cold agglutinin disease; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; GH01, General health item; Hb, Hemoglobin; IQR, Interquartile range; MWPC, Meaningful Within Patient Change; MCS, Mental Component Summary; PRO, Patient-reported outcome; PGI-S, Patient Global Impression of [fatigue] Severity; PGI-C, Patient Global Impression of Change; PCS, Physical Component Summary; RBC, Red blood cell; ROC, Receiver operating characteristic; SD, Standard deviation; SF-12v2, 12-Item Short-Form Health Survey - Version 2; US, United States.

Methods

Data sources

For qualitative analyses, patients with a confirmed diagnosis of CAD or cold agglutinin syndrome (previously known as secondary CAD) and their caregivers in the US were identified for inclusion in a survey panel. Quantitative analyses were based on data from two Phase 3 clinical trials evaluating sutimlimab as a treatment for CAD. Sutimlimab (ENJAYMOTM) is the only approved treatment for patients with CAD; it was first approved in 2022 in the USA (9). In the CARDINAL trial (NCT03347396), 24 adult patients with CAD with a recent history of blood transfusion and $\text{Hb} \leq 10.0$ g/dL received sutimlimab for 6 months during Part A of the trial (3). In the CADENZA trial (NCT03347422), adult patients with CAD without a recent history of blood transfusion and $\text{Hb} \leq 10.0$ g/dL were randomized 1:1 to receive sutimlimab (22 patients) or placebo (20 patients) for 6 months during Part A of the trial (10). The primary efficacy endpoint in both trials was a composite that required a Hb increase compared with baseline, no blood transfusions received, and no forbidden medications received. Fatigue was assessed as a secondary efficacy endpoint in both these trials using the FACIT-Fatigue questionnaire. In addition, other patient-reported outcome (PRO) tools, such as the 12-Item Short-Form Health Survey - Version 2 (SF-12v2[®]), the Patient Global Impression of [fatigue] Severity (PGI-S), and the Patient Global Impression of Change (PGI-C), provided exploratory efficacy endpoints.

Data from CARDINAL and CADENZA were pooled ($N = 66$) for a better CAD population representativity given the small sample size of both trials and the different inclusion criteria in terms of recent history of blood transfusions.

Instruments

FACIT-Fatigue

The FACIT questionnaires measure a range of health-related quality-of-life concepts in patients with cancer and other chronic diseases (11). FACIT-Fatigue is a 13-item questionnaire formatted

for self-administration. It assesses patient-reported fatigue and its impact on patient’s daily activities and functioning over the past 7 days (<https://www.facit.org/measures/FACIT-Fatigue>). Patients are specifically asked to answer each of the questions using a five-point Likert-type scale (0 = not at all; 1 = a little bit; 2 = somewhat; 3 = quite a bit; and 4 = very much). The total FACIT-Fatigue score ranges from 0 to 52, where higher scores represent less fatigue. The FACIT-Fatigue scale assesses the experience and impact of fatigue in a single global domain represented by the overall score (Figure 1) (12–14).

Other instruments

The SF-12v2 is a multipurpose, short-form health survey with 12 questions categorized into eight domains (subscales) of functioning and well-being, including two psychometrically based Physical Component Summary (PCS) and Mental Component Summary (MCS) measures. Each domain is scored on a T-score metric, which has a mean of 50 and a standard deviation of 10 (15). Higher scores (and positive change scores) on SF-12v2, PCS, and MCS indicate better daily functioning of patients. General health item (GH01) is the general health domain including only one question: *In general, would you say your health is: Excellent, Very Good, Good, Fair, Poor?*

The PGI-S is a self-reported, one-item questionnaire designed to assess participant’s impression of fatigue severity over the past 7 days using a five-point scale ranging from 1 to 5 (where 1 = none; 2

= mild; 3 = moderate; 4 = severe; and 5 = very severe). Higher scores indicate more severe fatigue.

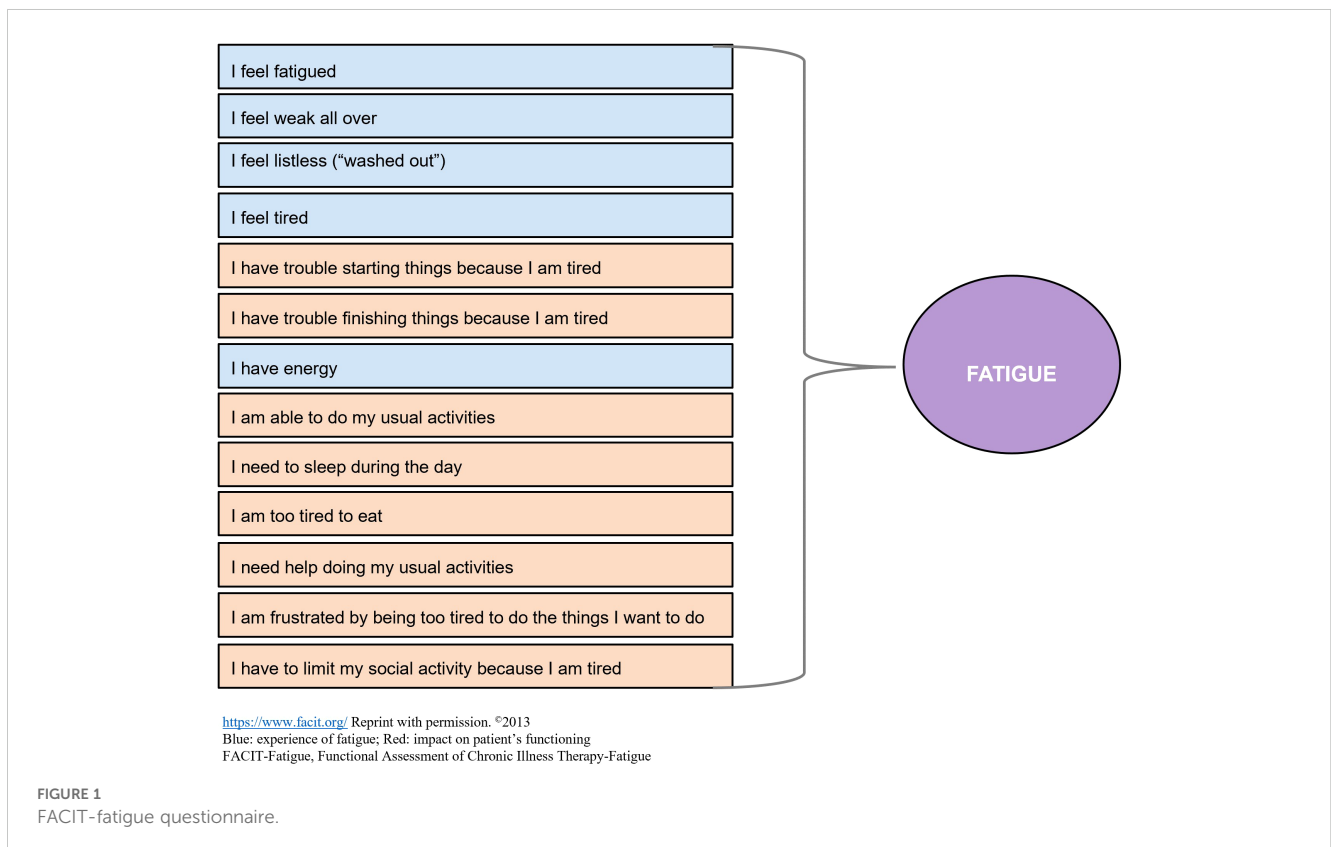
The PGI-C measures self-reported overall health change from baseline in a seven-point, single-item scale ranging from “very much worse” to “very much improved.”

Approach

Conceptual framework

The conceptual framework of the FACIT-Fatigue is characterized by having each of the 13 items contributing equally to the scoring of a single conceptual domain representing the “experience of fatigue” (blue coded box in Figure 1) and its “impact on patient’s functioning” (red coded box in Figure 1). A survey panel was conducted to identify the symptoms and impacts of CAD that are most important to patients with a confirmed diagnosis of CAD and their caregivers in the US, and to better understand their experience with CAD and its treatment. The sample was divided so that some participants were invited to the concept elicitation interviews and others to subsequent cognitive debriefing interviews. Furthermore, patients and some of their caregivers were interviewed to ascertain their understanding of the 13 items of the FACIT-Fatigue scale and whether these items were relevant to the CAD condition.

To test whether the measurement model of the FACIT-Fatigue was consistent with the hypothesized structure in patients with CAD, confirmatory factor analysis appropriate for categorical data was conducted using the pooled analytic sample from both studies.



Reliability

The internal consistency reliability of FACIT-Fatigue scores was evaluated using Cronbach's alpha coefficient. Cronbach's alpha was computed at baseline and Week 26 using listwise deletion of a single item at a time. The test-retest reliability of the FACIT-Fatigue was not investigated since neither CARDINAL nor CADENZA was designed to evaluate the stability of scores over time.

Convergent validity

Convergent validity was examined by evaluating the direction and magnitude of correlations between FACIT-Fatigue scores and scores from instruments measuring conceptually related PROs or clinical measures. Several such instruments were used to evaluate the convergent validity of FACIT-Fatigue scores, including SF-12v2, PCS and MCS, PGI-S, and Hb level. Convergent validity between FACIT-Fatigue scores and each of the instrument measures was calculated at baseline and Week 26 using Spearman rank-order correlation coefficient, where both FACIT-Fatigue and the respective instruments were concurrently assessed. Correlations of ≥ 0.40 were interpreted as strong evidence for convergent validity. Correlations of 0.20–0.39 were considered to provide more limited support for convergent validity.

Known-group validity

Known-group validity tests whether a PRO scale can discriminate between groups of patients who differ on a conceptually related criterion measure in the direction and magnitude that is consistent with a prespecified hypothesis (16). Hb and PGI-S were used to assess known-group validity of the FACIT-Fatigue scale in CAD.

Ability to detect change

The evidence to support the ability of FACIT-Fatigue to detect change was produced in two ways. The responsiveness of FACIT-Fatigue was investigated using Spearman rank-order correlations between changes in FACIT-Fatigue scores and changes in SF-12v2 PCS and MCS scores, changes in PGI-S scores, and changes in Hb levels from baseline to Week 26 (or to turn around time for Hb). FACIT-Fatigue's ability to detect change was also evaluated by comparing mean changes in scores between groups of patients with different Hb levels and between groups of patients with different PGI-C responses.

Meaningful Within-Patient Change

The Meaningful Within Patient Change (MWPC) reflects a score change in a PRO instrument that indicates a meaningful

treatment benefit for individual patients. Anchor-based and distribution-based methods were used to estimate the MWPC of the FACIT-Fatigue in patients with CAD. Anchors used in this analysis were: the GH01 of SF-12v2, PGI-S, PGI-C, and Hb level (g/dL). Three different anchor-based analyses were applied: mean score change, receiver operating characteristic (ROC) curve (model-based adjusted), and logistic regression (model-based). A detailed description is available in the [Supplementary Material](#).

Results

Qualitative analyses results

For the qualitative study, 22 patients with CAD expressed interest in participating in the concept elicitation interviews. Out of the 22 patients, 18 were contacted and screened, and 16 patients were found eligible. The 16 patients were white, and their mean age was 67.8 years. Most of these patients were female (62.5%) with three-fourths having a college or professional degree. The patient population was geographically diverse, with half of the patients residing in the South or Midwest of the USA. The average time since CAD diagnosis was approximately 6 years; 68.8% of these patients had been diagnosed with CAD within the last 1.5 years. Fatigue was the most common and bothersome symptom associated with CAD in the concept elicitation interviews and had a significant impact on a patient's ability to perform routine daily activities. All patients, except one (15/16; 93.8%), reported feeling fatigued due to CAD. Among these, 12 patients (80%) experienced fatigue daily, while three patients (20%) reported feeling fatigued few days per week. Overall, 10 (66.7%) of these 15 patients described their fatigue/tiredness/lack of energy as physical compared with two patients (13.3%) who described it as mental and three patients (20%) who described it as both physical and mental.

Outcomes for FACIT-Fatigue

Following the concept elicitation portion of the interview, participants reviewed and provided feedback on the FACIT-Fatigue scale, which was sent to each participant before their scheduled interview. The impacts identified by patients in the concept elicitation interviews mapped well with the item content measuring fatigue impact in the FACIT-Fatigue questionnaire. Overall, patients were able to understand and interpret the FACIT-Fatigue scale instructions, items, response options, and recall period without any problem. Nearly all patients ($n = 15$; 93.8%) reported that it was easy for them to recall the past 7 days. Each item of the FACIT-Fatigue questionnaire was reported to be clear and easy to understand by most of the patients ($n = 15$; 93.8%). All patients could answer all 13 items; however, one patient reported that she was unsure what "listless" or "washed out" meant.

All patients ($n = 16$; 100%) reported that the FACIT-Fatigue questionnaire could capture their experiences with CAD-related fatigue. All patients ($n = 16$; 100%) reported that FACIT-Fatigue

measured a concept important to them and was not missing any important elements. However, none of the items were identified as most important by the majority of patients (Supplementary Table 1). Two items “I feel fatigued” and “I have trouble starting things because I’m tired” were the most frequently selected items and considered most important to patients. In contrast, three items were not identified by any of the patients as being important to them: “I am too tired to eat,” “I need help doing my usual activities,” and “I need to sleep during the day” (Supplementary Table 1).

The cognitive debriefing interviews that were conducted with an additional set of 21 patients with CAD confirmed that the items of the FACIT-Fatigue questionnaire were easy to understand and relevant to the experience of patients with CAD.

Quantitative analyses results

Baseline characteristics

Overall, 66 patients were enrolled in the CARDINAL ($n = 24$) and CADENZA ($n = 42$) trials. Of these, 55 patients had FACIT-Fatigue scores available at baseline and Week 26 and were included in this pooled analytic sample. Most of the included patients were female (76.4%); mean (standard deviation, SD) age was 68.8 (10.1) years and mean (SD) FACIT-Fatigue score was 32.6 (11.2) at baseline (Table 1).

Results from the CFA showed that item-to-factor loadings were acceptable (Supplementary Table 2).

Reliability of FACIT-Fatigue in CAD

The internal consistency reliability of the FACIT-Fatigue scale far exceeded the minimum threshold for acceptable reliability. Cronbach’s alpha coefficient was 0.94 at baseline and 0.96 at Week 26. Deletion of one item at a time did not change the magnitude of Cronbach’s alpha coefficient.

Convergent validity

Data from the pooled analytic sample reported a correlation of 0.24 between FACIT-Fatigue scores and Hb levels at baseline and that of 0.39 at Week 26 (Table 2). The lower correlation at baseline reflected the truncated distribution of Hb levels due to the trial inclusion criteria ($Hb \leq 10.0$ g/dL in both CARDINAL and CADENZA studies).

FACIT-Fatigue scores were highly correlated with the PCS and MCS measures of SF-12v2 and PGI-S at baseline and Week 26. All correlations exceeded the recommended (absolute value) level of 0.40 for “strong” convergent validity (Table 2).

Known-group validity

Data from the pooled analytic sample were utilized for known-group validity tests of the FACIT-Fatigue scale in

patients with CAD at baseline and Week 26 based on Hb levels and PGI-S outcomes. The mean FACIT-Fatigue score was significantly higher (indicating less fatigue) among patients in the higher Hb level group (≥ 12 g/dL) than the lower Hb level group (< 12 g/dL) at Week 26, while it was numerically higher in patients in the ≥ 9 g/dL Hb group than the ≤ 9 g/dL Hb group at baseline (Table 3).

For PGI-S at baseline, patients reporting “severe” fatigue had the lowest FACIT-Fatigue score, mean (SD): 16.8 (3.4), followed by patients reporting “moderate” fatigue severity: 30.6 (10.3). Patients reporting “none/mild” fatigue severity had the highest FACIT-Fatigue score – none: 47.7 (3.6); mild: 37.5 (5.7). Similarly, mean FACIT-Fatigue scores differed significantly ($P < 0.001$) across patients who differed in the patient global assessment of fatigue severity at Week 26. Patients reporting “severe” fatigue had the lowest FACIT-Fatigue score, mean (SD): 19.2 (8.2), followed by patients reporting “moderate” fatigue severity: 35.7 (6.2). Patients reporting “none/mild” fatigue severity had the highest FACIT-Fatigue score – none: 47.3 (6.4); mild: 45.3 (3.2) (Table 3).

Ability to detect change

Changes in FACIT-Fatigue scores were moderately correlated with changes in Hb levels ($r = 0.52$, $P < 0.001$), SF-12v2 PCS ($r = 0.45$, $P < 0.001$), and MCS ($r = 0.59$, $P < 0.001$) measures, and PGI-S ($r = -0.76$, $P < 0.001$). The differences in mean changes in FACIT-Fatigue scores between Hb responders and non-responders were large and statistically significant (11.3 vs. 1.5, $P = 0.0045$) (Table 4). Mean FACIT-Fatigue score changes were also statistically significantly different across categories of the PGI-C ($P = 0.0027$) (Table 4). The mean improvement in FACIT-Fatigue scores was largest among patient groups reporting the largest improvement in PGI-C, while patients who reported that their PGI-C was the same as or worse than baseline showed little change or worsening in FACIT-Fatigue scores.

Meaningful within-patient change

Anchor-based estimates: Polychoric correlations between FACIT-Fatigue scores and each of the four anchors (GH01 of SF-12v2, PGI-S, PGI-C, and Hb level) were computed at baseline, at Week 26, and for the change in score (from baseline to Week 26). These correlations mostly exceeded 0.40, indicating moderate-to-strong associations and supporting the use of these anchors to estimate the MWPC for FACIT-Fatigue in CAD (Supplementary Table 3).

MWPC thresholds for FACIT-Fatigue in CAD, as estimated from the anchor-based methods, are presented in Table 5. In the mean change approach, the subset of patients ranged between 10 and 30 across the four anchors. The lowest mean (SD) change estimate in FACIT-Fatigue score was 9.2 (16.0) for PGI-C, and the highest was 15.7 (10.6) for PGI-S. The analyses of mean change produced the highest MWPC estimates of any of the analyses. In the ROC curve approach, the MWPC estimates ranged from 2.0 for Hb level to 8.0

TABLE 1 Patient demographics and clinical characteristics in the pooled analytic sample.

Patient demographics	CARDINAL and CADENZA analytical samples	
	N	%
Trial (n = 55)		
CADENZA	38	69.09
CARDINAL	17	30.91
Treatment received (n = 55)		
Sutimlimab ^a IV infusion	36	65.45
Placebo	19	34.55
Gender (n = 55)		
Female	42	76.36
Male	13	23.64
	Mean ± SD	Median (min, max)
Age (n = 55)	68.8 ± 10.10	70.0 (46.0, 88.0)
PRO measures ^b	Mean ± SD	Median (min, max)
FACIT-Fatigue		
FACIT-Fatigue at baseline (n = 55)	32.6 ± 11.2	33.6 (9.0, 51.0)
FACIT-Fatigue at Week 26 (n = 55)	39.3 ± 10.9	42.0 (8.0, 52.0)
Change in FACIT-Fatigue (n = 55)	6.7 ± 13.1	4.0 (-24.0, 38.0)
GH01		
GH01 at baseline (n = 53)	3.5 ± 0.7	3.0 (2.0, 5.0)
GH01 at Week 26 (n = 54)	3.1 ± 0.9	3.0 (1.0, 5.0)
Change in GH01 (n = 53)	-0.3 ± 0.8	0.0 (-3.0, 1.0)
PGIS^c		
PGIS at baseline (n = 35)	2.5 ± 1.0	3.0 (1.0, 4.0)
PGIS at Week 26 (n = 35)	2.2 ± 1.0	2.0 (1.0, 4.0)
Change in PGIS (n = 35)	-0.3 ± 1.0	0.0 (-2.0, 2.0)
PGIC		
PGIC at Week 26 (n = 53)	3.0 ± 1.4	3.0 (1.0, 6.0)
	N	%
Categories of PGIC at Week 26 (n = 53)		
Much worse	3	5.66
Minimally worse	4	7.55
No change	12	22.64
Minimally improved	10	18.87
Much improved	18	33.96
Very much improved	6	11.32
Clinical/laboratory measures	Mean ± SD	Median (min, max)

(Continued)

TABLE 1 Continued

Patient demographics	CARDINAL and CADENZA analytical samples	
	N	%
Hb		
Hb at baseline (n = 55)	9.05 ± 1.16	9.10 (4.90, 11.20)
Hb at average of Weeks 23, 25, and 26 (n = 55)	10.81 ± 2.06	11.03 (5.93, 13.87)
Change in Hb (n = 55)	1.76 ± 2.00	2.23 (-2.47, 6.20)

^aFormerly BIVV009.
^bHigher scores (and positive change scores) on FACIT-Fatigue indicate better functioning. Lower scores (and negative change scores) on PGI-S and PGI-C indicate better functioning.
^cThe recommended anchor, PGI-S, was implemented after the start of the CARDINAL trial; therefore, not all patients had PGI-S scores available at baseline.
 FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; GH01, general health item of SF-12v2; Hb, hemoglobin; IV, intravenous; PGI-C, Patient Global Impression of Change; PGI-S, Patient Global Impression of [fatigue] Severity; PRO, patient-reported outcome; SD, standard deviation; SF-12v2, 12-Item Short-Form Health Survey - Version 2.

for GH01. In the logistic regression approach, the MWPC estimates were highest for GH01 (8.9) and lowest for PGI-C (4.2).

Distribution-based estimates: Using effect size, the MWPC based on one-half of SD of FACIT-Fatigue at baseline was 5.6. The MWPC based on the standard error of measurement, which incorporates the SD and the internal consistency reliability of the FACIT-Fatigue at baseline, was estimated to be 2.8.

Retained MWPC: FACIT-Fatigue MWPC estimates produced from the anchor-based methods ranged from 2.0 to 15.7 (interquartile range [IQR]: 4.8–9.7) with a median of 7.5. The mean score change analyses were highly sensitive to distributional characteristics of data (outliers) given that only a subset of the patient data was used (the most significant limitation of the mean change method). When considering the ROC and regression-based analyses only, the IQR of MWPC estimates was 4.1–7.3, and individual responder definitions were in the range of 5–8 points, where “5” is the lowest recommended MWPC threshold for FACIT-Fatigue in CAD, and “8” is the most stringent threshold.

TABLE 2 Convergent validity assessment of FACIT-Fatigue in the pooled analytic sample.

Outcome measure	Correlation with FACIT-Fatigue score	
	Baseline	Week 26
Hb	r = 0.24	r = 0.39
SF-12v2 PCS	r = 0.62	r = 0.64
SF-12v2 MCS	r = 0.47	r = 0.69
PGI-S	r = -0.75	r = -0.78

FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; Hb, hemoglobin; MCS, Mental Component Summary; PCS, Physical Component Summary; PGI-S, Patient Global Impression of [fatigue] Severity; SF-12v2, 12-Item Short-Form Health Survey - Version.

Convergent validity between FACIT-Fatigue scores and each of the instrument measures was calculated at baseline and Week 26 using Spearman rank-order correlation coefficient, where both FACIT-Fatigue and the respective instruments were concurrently assessed. Correlations of ≥0.40 were interpreted as strong evidence for convergent validity. Correlations of 0.20–0.39 were considered to provide more limited support for convergent validity.

TABLE 3 Known-group validity assessment of the FACIT-Fatigue scale in relation to Hb and PGI-S in the pooled analytic sample.

Outcome measure	Comparison group	FACIT-Fatigue score at baseline, mean (SD)	FACIT-Fatigue score at Week 26, mean (SD)	P-value
Hb	Hb ≤ 9 g/dL (n = 27)	29.75 (12.2)	-	0.0691
	Hb ≥ 9 g/dL (n = 28)	35.26 (9.8)	-	
	Hb ≥ 12 g/dL (n = 20)	-	43.3 (9.0)	<0.05
	Hb <12 g/dL (n = 35)	-	36.9 (11.3)	
PGI-S ^{a,b}	None	47.7 (3.6)	47.3 (6.4)	At baseline: <0.001 At Week 26: • <0.001
	Mild	37.5 (5.7)	45.3 (3.2)	
	Moderate	30.6 (10.3)	35.7 (6.2)	
	Severe	16.8 (3.4)	19.2 (8.2)	

^aN values differed for PGI-S severity groups at baseline and Week 26. At baseline: none, n = 6; mild, n = 10; moderate, n = 14; and severe, n = 5. At Week 26: none, n = 9; mild, n = 14; moderate, n = 17; and severe, n = 5.

^bThere were no patients in the CARDINAL and CADENZA studies who had the PGI-S “very severe” outcome – neither at baseline nor at Week 26.

The Hb groups differed at baseline (≤9 g/dL and ≥9 g/dL) and Week 26 (≤12 g/dL and ≥12 g/dL) as Hb levels were part of the screening criteria of CARDINAL and CADENZA studies. Mean FACIT-Fatigue scores were compared between the two hemoglobin groups using Student’s t-test.

For PGIS, the statistical significance of differences in mean FACIT-Fatigue scores was tested using analysis of variance methods (ANOVA).

FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; Hb, hemoglobin; PGI-S, Patient Global Impression of [fatigue] Severity; SD, standard deviation.

TABLE 4 Ability of FACIT-Fatigue to detect change in PRO and clinical outcome measures in pooled analytic sample.

Response group	N	Mean change (SD) FACIT-Fatigue	P-value
Hb responder (change ≥2 g/dL)	29	11.3 (12.0)	0.0045
Hb non-responder (change <2 g/dL)	26	1.5 (12.4)	
PGI-C “very much improved”	6	16.3 (11.3)	0.0027
PGI-C “much improved”	18	10.8 (11.5)	
PGI-C “minimally improved”	10	9.2 (16.0)	
PGI-C “no change”	12	2.9 (5.6)	
PGI-C “minimally worse”	4	-3.4 (11.0)	
PGI-C “much worse”	3	-16.5 (6.5)	

There were no patients in the PGI-C “very much worse” category (n = 0).

Mean FACIT-Fatigue scores were compared between the two hemoglobin groups using Student’s t-test.

For PGI-C, the statistical significance of differences was tested using analysis of variance methods (ANOVA).

FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; Hb, hemoglobin; PGI-C, Patient Global Impression of Change; PRO, patient-reported outcome; SD, standard deviation.

Discussion

Fatigue is a key symptom among patients with CAD. The various analyses conducted in this study were aimed to assess the validity of using FACIT-Fatigue and to estimate the related MWPC in patients with CAD. Concept elicitation interviews were conducted with patients with CAD in the USA and allowed to qualitatively evaluate the content validity of the FACIT-Fatigue scale in CAD. Concept mapping showed that elicitation interviews mapped well with the item content of FACIT-Fatigue measuring fatigue impact in patients with CAD. Results from the cognitive

debriefing interviews reported FACIT-Fatigue as an overall relevant scale in CAD, and patients could correctly interpret each item and felt the items to be relevant to their experiences.

Post hoc analyses using the pooled analytic sample from sutimlimab Phase 3 trials were also conducted to quantitatively assess the psychometric properties of the FACIT-Fatigue scale in CAD. Overall, the evidence produced in this quantitative analysis supports the use of FACIT-Fatigue in patients with CAD. Estimates of internal consistency reliability far exceeded the minimum standard of 0.70. The FACIT-Fatigue scale showed strong convergence with other relevant scales, including the PCS and MCS measures of SF-12v2 and PGI-S. Correlation with baseline Hb levels was lower (0.24), but that may be a result of the truncated distribution of Hb levels due to trial inclusion criteria (required Hb level ≤10 g/dL at screening). Tests of known-group validity based on Hb and PGI-S were statistically significant and supported the discriminant validity of FACIT-Fatigue scores in patients with CAD. Similarly, tests of the responsiveness based on Hb and PGI-C were statistically significant and strongly supported the validity of the FACIT-Fatigue in detecting change in patients with CAD. Results from this analysis demonstrated that meaningful reductions in fatigue were associated with improvement in Hb levels – a main clinical outcome measure of CAD. There were large differences in MWPC estimates derived using the mean change method compared with other anchor-based methods. The mean change method was subject to considerable influence from outliers due to the smaller sample size (subgroup of patients) producing larger estimates than other methods. Hence, when estimating the MWPC for FACIT-Fatigue, ROC and regression-based analyses were used, producing the IQR of 4.1–7.3 for MWPC estimates. The corresponding response criteria ranged from 5–8 points, where “5” is the lowest recommended MWPC threshold for FACIT-Fatigue in CAD.

The clinically meaningful change analyses for FACIT fatigue in patients with CAD were previously presented at the 26th Congress of the European Hematology Association, with a reported threshold

TABLE 5 Anchor-based analyses – results of the mean change, ROC-based, and logistic-regression-based analyses for FACIT-Fatigue score.

Anchors	Total, N	Improved, N (%)	Mean change analyses		ROC-based analyses				Logistic-regression-based analyses
			Responders ^a N	Mean (SD) change in FACIT-Fatigue score	AUC	Sensitivity	Specificity	MWPC	Adjusted MWPC
GH01 (SF-12v2)	53	21 (39.6)	18	14.4 (12.7)	0.79	0.71	0.78	8.0	8.9
PGI-S	35	15 (42.9)	11	15.7 (10.6)	0.93	0.87	0.85	5.0	7.0
PGI-C	53	34 (64.2)	10	9.2 (16.0)	0.74	0.68	0.74	4.0	4.2
Hb	55	30 (54.5)	30	11.1 (11.9)	0.73	0.73	0.68	2.0	5.7

For each model, an anchor-based binary indicator of change served as the dependent variable and the FACIT-F change score served as the independent variable. ^aResponders were defined as follows: for GH01 and PGI-S, patients who had a one-point improvement between Week 26 and baseline; for PGI-C, patients who selected a response of “minimally improved”; and for Hb, all responders based on the CARDINAL and CADENZA criteria (respectively Hb change ≥ 2g/dL and Hb change ≥ 1.5g/dL). AUC, area under receiver operating characteristic curve; FACIT-Fatigue, Functional Assessment of Chronic Illness Therapy-Fatigue; GH01, general health item of SF-12v2; Hb, hemoglobin; MWPC, meaningful within-patient change; PGI-C, Patient Global Impression of Change; PGI-S, Patient Global Impression of [fatigue] Severity; ROC, receiver operating characteristic; SD, standard deviation; SF-12v2, 12-Item Short-Form Health Survey – Version.

of 5 (17). Several analyses have estimated the clinically meaningful thresholds of FACIT-Fatigue score in other chronic diseases with larger sample sizes of patients (18–22). In these studies, the meaningful threshold estimates ranged from 3 to 10 points; hence, the MWPC estimate for patients with CAD from the current analysis is in line with these previous findings.

Limitations

The CARDINAL and CADENZA trials had small patient numbers. To mitigate the impact of this, and the single-arm design of the CARDINAL trial, data were pooled from both trials. Nevertheless, the pooled sample size remained small with regard to the reported analyses, in particular the fit statistics. Although three of the four pre-specified fit statistics suggested uni-dimensionality, one (RMSEA) did not meet the prespecified criterion. Traditional cutoffs and standards for CFA fit statistics are not universally recommended when evaluating dimensionality. Cook et al. concluded that an investigative approach such as this is favored over reliance on published criteria, and that bifactor analysis is appealing when one wishes to account for evaluation of the relative impact of secondary dimensions such as experience and impact. These results would therefore benefit from replication in a larger sample (23).

The *post hoc* nature of these analyses remains a limitation.

Conclusion

Results from the analyses presented above show that FACIT-Fatigue is a reliable, valid, responsive PRO measure for patients with CAD. MWPC estimates obtained in this study for the FACIT-Fatigue questionnaire in patients with CAD were consistent with those published previously. These analyses confirm the validity of using the FACIT-Fatigue scale in CAD and provide a clinically meaningful threshold for future CAD studies.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://ashpublications.org/blood/article/140/9/980/485526/> Sutinlimab-in-patients-with-cold-agglutin in https://www.nejm.org/doi/10.1056/NEJMoa2027760?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed.

Ethics statement

The studies involving humans were approved by RTI Int. Internal Review Board. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants’ legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

DC: Conceptualization, Data curation, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. QH: Conceptualization, Data curation, Formal Analysis, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. AK: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing. FJ: Conceptualization, Data curation, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing. TS: Writing – review & editing. FS: Writing – review & editing. MW: Writing – review & editing. MK: Conceptualization, Data curation, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. AR: Conceptualization, Data curation, Investigation, Methodology, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This study was funded by Sanofi. The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article, or the decision to submit it for publication.

Acknowledgments

Scientific writing support for this manuscript, under the guidance of the authors, was provided by Deepti Sharda and was funded by Sanofi in accordance with Good Publication Practice guidelines.

Conflict of interest

DC received research funding from Alexion, Akcea, Novartis, Pfizer, and Merck and consultancy fees from Alexion, Apellis, Daiichi Sankyo, Novartis, Pfizer, and Sanofi. QH received honoraria for lecturing or advisory work from Alexion, Amgen, Apellis, argenx, Grifols, Incyte, Immunovant, Janssen, Novartis, ReAlta, Sanofi, and Sobi. AK is a consultant for Sanofi and was

employed by BMAPS SARM. FJ, TS, FS, and MW are employees of Sanofi and may own stock options in the company. MK is an employee of QualityMetric, a consulting company that received funds from Sanofi to conduct this study. AR is a consultant for and received honoraria from Alexion, Amgen, Apellis, BioCryst, Grifols, Kira, Novartis, Roche, Sanofi, Sobi, and Bioverativ and also received research funding from Roche.

The reviewer SB declared a past co-authorship with the authors FS, MW, and AR to the handling editor.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/frhem.2025.1490130/full#supplementary-material>

References

- Gard N. Cold agglutinin disease 2021. Available online at: <https://rarediseases.info.nih.gov/diseases/6130/cold-agglutinin-disease>. (Accessed February 17, 2025).
- Mullins M, Jiang X, Bylsma LC, Fryzek JP, Reichert H, Chen EC, et al. Cold agglutinin disease burden: a longitudinal analysis of anemia, medications, transfusions, and health care utilization. *Blood Adv.* (2017) 1:839–48. doi: 10.1182/bloodadvances.2017004390
- Röth A, Barcellini W, D'Sa S, Miyakawa Y, Broome CM, Michel M, et al. Sutimlimab in cold agglutinin disease. *N Engl J Med.* (2021) 384:1323–34. doi: 10.1056/NEJMoa2027760
- Berentsen S, Barcellini W, D'Sa S, Randen U, Tvedt THA, Fattizzo B, et al. Cold agglutinin disease revisited: a multinational, observational study of 232 patients. *Blood.* (2020) 136:480–8. doi: 10.1182/blood.2020005674
- Hansen DL, Möller S, Frederiksen H. Survival in autoimmune hemolytic anemia remains poor, results from a nationwide cohort with 37 years of follow-up. *Eur J Haematol.* (2022) 109:10–20. doi: 10.1111/ejh.v109.1
- Dhaliwal G, Cornett PA, Tierney LM Jr. Hemolytic anemia. *Am Fam Physician.* (2004) 69:2599–606.
- Jun Su KK, Dana diBenedetti. Patient-reported disease burden: in-depth interviews of patients with CAD. *Blood.* (2020) 136:29–30. doi: 10.1182/blood-2020-136788
- Joly F, Schmitt LA, Watson PAM, Pain E, Testa D. The burden of cold agglutinin disease on patients' Daily life: web-based cross-sectional survey of 50 american patients. *JMIR Form Res.* (2022) 6:e34248. doi: 10.2196/34248
- FDA. FDA approves treatment for adults with rare type of anemia 2022. Available online at: <https://www.fda.gov/drugs/news-events-human-drugs/fda-approves-treatment-adults-rare-type-anemia>. (Accessed February 17, 2025).
- Röth A, Berentsen S, Barcellini W, D'Sa S, Jilma B, Michel M, et al. Sutimlimab in patients with cold agglutinin disease: results of the randomized placebo-controlled phase 3 CADENZA trial. *Blood.* (2022) 140:980–91. doi: 10.1182/blood.2021014955
- Webster K, Cella D, Yost K. The Functional Assessment of Chronic Illness Therapy (FACIT) Measurement System: properties, applications, and interpretation. *Health Qual Life Outcomes.* (2003) 1:79. doi: 10.1186/1477-7525-1-79
- Cella D, Lai JS, Stone A. Self-reported fatigue: one dimension or more? Lessons from the Functional Assessment of Chronic Illness Therapy–Fatigue (FACIT-F) questionnaire. *Support Care Cancer.* (2011) 19:1441–50. doi: 10.1007/s00520-010-0971-1
- Chandran V, Bhella S, Schentag C, Gladman DD. Functional assessment of chronic illness therapy-fatigue scale is valid in patients with psoriatic arthritis. *Ann Rheum Dis.* (2007) 66:936–9. doi: 10.1136/ard.2006.065763
- Pouchot J, Kherani RB, Brant R, Laccaille D, Lehman AJ, Ensworth S, et al. Determination of the minimal clinically important difference for seven fatigue measures in rheumatoid arthritis. *J Clin Epidemiol.* (2008) 61:705–13. doi: 10.1016/j.jclinepi.2007.08.016
- Ware J Jr., Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* (1996) 34:220–33. doi: 10.1097/00005650-199603000-00003
- Kerlinger. *Foundations of Behavioral Research.* New York: Second Holt, Rinehart and Winston (1973).
- Hill QA, Röth A, Jilma B, Broome CM, Berentsen S, Rizio AA, et al. Clinically important change in facit-fatigue score for patients with cold agglutinin disease: An analysis using the phase 3 cardinal and cadenza studies. New York: EHA2021 virtual congress abstract book. *HemaSphere.* (2021) 5:e566.
- Cella D, Yount S, Sorensen M, Chartash E, Sengupta N, Grober J. Validation of the Functional Assessment of Chronic Illness Therapy Fatigue Scale relative to other instrumentation in patients with rheumatoid arthritis. *J Rheumatol.* (2005) 32:811–9.
- Lai JS, Beaumont JL, Ogale S, Brunetta P, Cella D. Validation of the functional assessment of chronic illness therapy-fatigue scale in patients with moderately to severely active systemic lupus erythematosus, participating in a clinical trial. *J Rheumatol.* (2011) 38:672–9. doi: 10.3899/jrheum.100799
- Reddy S, Bruera E, Pace E, Zhang K, Reyes-Gibby CC. Clinically important improvement in the intensity of fatigue in patients with advanced cancer. *J Palliative Med.* (2007) 10:1068–75. doi: 10.1089/jpm.2007.0007
- Cella D, Eton DT, Lai JS, Peterman AH, Merkel DE. Combining anchor and distribution-based methods to derive minimal clinically important differences on the Functional Assessment of Cancer Therapy (FACT) anemia and fatigue scales. *J Pain symptom management.* (2002) 24:547–61. doi: 10.1016/S0885-3924(02)00529-8
- Goligher Ec PJ, Brant R, Kherani R, Aviña-Zubieta J, Laccaille D, Lehman A, et al. Minimal clinically important difference for 7 measures of fatigue in patients with systemic lupus erythematosus. *J Rheumatol.* (2008) 35:635–42.
- Cook KF, Kallen MA, Amtmann D. Having a fit: impact of number of items and distribution of data on traditional criteria for assessing IRT's unidimensionality assumption. *Qual Life Res.* (2009) 18:447–60. doi: 10.1007/s11136-009-9464-4