Susan (Sue) Vallow susan.vallow@novartis.com

Psychometric Performance of the Functional Assessment of Chronic Illness Therapy (FACIT) Fatigue Questionnaire Among Patients with Paroxysmal Nocturnal Hemoglobinuria

David Cella,[1] Susan Vallow,[2] Georgina Bermann, [3] Jeffrey McDonald,[4] Gilbert Ngerano,[4] Samantha Linton,[4] Ethan Arenson,[4] Samopriyo Maitra,[5] Roger Lamoureux,[4] Gavin Dickie[4]

[1]Northwestern University, Evanston, IL, USA; [2]Novartis Pharmaceutical Corporation, East Hanover, NJ, USA; [3]Novartis Pharma AG, Basel, Switzerland; [4]Adelphi Values, Boston, MA, USA; [5]Novartis Healthcare Pvt. Ltd., Hyderabad, Rangareddy, India

This study is sponsored by Novartis Pharma AG
Poster presented at ISPOR Europe 2023 held on 12–15 November 2023

Scan to obtain poster

https://www.medicalcongress.novartisoncology.com/ISPOREU23/PNH/pnh.html#SusanVallowEU1

KEY FINDINGS & CONCLUSIONS

- Testing for essential unidimensionality of the FACIT-Fatigue supported the reporting of the FACIT-Fatigue total score in the APPLY-PNH and APPOINT-PNH clinical trials
- Across both trials, the FACIT-Fatigue was found to be reliable, with high internal consistency and strong test-retest reliability for the total score
- FACIT-Fatigue total scores were moderately to strongly correlated with scores of supplementary assessments
- Changes in FACIT-Fatigue total score were strongly correlated with changes in EORTC-QLQ-C30 fatigue score and the patient global impression of fatigue severity score, and weakly correlated with changes in hemoglobin level
- The FACIT-Fatigue total score was able to distinguish between distinct groups defined by patient global impression of fatigue severity scores and hemoglobin change
- Overall, these results demonstrate that the FACIT-Fatigue total score is appropriate for measuring change in fatigue in the context of these clinical trials

INTRODUCTION

- Paroxysmal nocturnal hemoglobinuria (PNH) is a rare hemolytic disorder characterized by hemolysis and bone marrow failure leading to anemia. Fatigue is a key and frequently-reported symptom of PNH
- The analyses reported here evaluated the psychometric properties of the FACIT-Fatigue questionnaire in two Phase 3 clinical trials of iptacopan, a novel oral treatment for PNH
  - APPLY-PNH was a randomized, active-controlled, open-label trial
  - APPOINT-PNH was a single-arm, open-label trial
- A key secondary objective for both clinical trials included evaluation of the effect of iptacopan on improving fatigue related to PNH, using the FACIT-Fatigue total score

METHODS

- Factor structure of the FACIT-Fatigue was evaluated for essential unidimensionality to support reporting of the FACIT-Fatigue total score
- Reliability analyses assessed the internal consistency and stability over time of the FACIT-Fatigue total score
- Concurrent validity of the FACIT-Fatigue total score was assessed by examining correlations with:
  - A question on the patient's global impression of fatigue severity (PGIS)
  - European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire – Core 30 (EORTC-QLQ-C30) scores
  - Five-level EQ-5D (EQ-5D-5L) scores
- Ability to distinguish between clinically distinct groups was assessed by PGIS level and hemoglobin change
- Responsiveness to change was assessed by anchoring FACIT-Fatigue total score to changes in EORTC-QLQ-C30 fatigue scores, PGIS scores, and hemoglobin levels

RESULTS

Analysis populations

- Analysis population for APPLY-PNH (N=95) was 68% male, median age 53.0 years (**Table 1)**
- Analysis population for APPOINT-PNH (N=40) was 57% male, median age 38.5 years (**Table 1**)
- Essential unidimensionality was evaluated using pooled data from both the APPLY-PNH and APPOINT-PNH analysis populations

Essential unidimensionality of the FACIT-Fatigue

- To support the appropriateness of the FACIT-Fatigue total score for measuring improvement in fatigue, the factor structure of the FACIT-Fatigue was evaluated for essential unidimensionality using a bifactor model[1,2]
- A bifactor model assumes that each item in the FACIT-Fatigue relates to the overall concept of fatigue, while allowing for respective grouping of items relating to the Symptom experience or Impacts domains
- Essential unidimensionality would be demonstrated demonstrated if:
  - The percentage of all correlations among items attributable purely to the overall concept of fatigue (percentage of uncontaminated correlation [PUC}) was ≥ 80%; OR
  - The PUC was <80% but there was a high proportion of items whose variance was explained by the overall concept of fatigue (percentage explained common variance [PECV]); AND
  - The mean relative bias was below 15% in magnitude[3]
- Analysis was conducted at Day 140 and Day 168 for the pooled APPLY-PNH and APPOINT-PNH analysis populations
- The mean relative bias estimates at both timepoints were well below the 15% threshold
- PUC was <80% at both timepoints, but PECV was high (98.2% at Day 140, and 97.1% at Day 168)[4] (**Table 2**)
- Results provide evidence of the essential unidimensionality of the FACIT-Fatigue and support the use and reporting of the FACIT-Fatigue total score

Reliability of the FACIT-Fatigue total score

- **Internal consistency** of the FACIT-Fatigue was investigated by calculating (1) Cronbach's alpha coefficient (α)[5] which is a lower bound to internal consistency, and (2) McDonald's omega coefficient (ω)[6] (**Table 3)**
  - FACIT-Fatigue total score demonstrated very high internal consistency at Baseline, Days 42, 126, 140, and 168 with estimates of Cronbach's α and McDonald's ω above 0.90
- **Test-retest reliability** analyses were conducted at Screening and Day 1 (**Table 4**), computed as intra-class coefficient (ICC) for absolute agreement in a two-way mixed effects model
  - Upper bound of the 95% confidence interval for both trials was above 0.90, suggesting very good test-retest reliability for the FACIT-Fatigue total score

Concurrent validity of the FACIT-Fatigue total score

- Correlations were computed for each supplementary assessment at Baseline, Day 42, Day 126, Day 140, and Day 168 in each trial (**Table 5** shows range of correlations across all timepoints)
- Correlations of at least |0.30|[7] were considered sufficient evidence of concurrent validity
- Almost all supplementary assessments were moderately to strongly correlated with the FACIT-Fatigue total score across all timepoints for both trials

Responsiveness to change of the FACIT-Fatigue total score

- Mean change score for changes between FACIT-Fatigue total score and its associated Cohen's *d* effect size[8] statistic were computed between Baseline and Day 42 and between Baseline and Day 168
- Next, the strength of correlation between change in supplementary assessment scores and change in FACIT-Fatigue total scores were examined (**Table 6)**
- Change estimates were not interpreted by statistical significance but by magnitude of change
- Changes in the EORTC-QLQ-C30 fatigue score and the PGIS score had strong to very strong correlations to changes in the FACIT-Fatigue total score at both Day 42 and Day 168 for both trials
- As expected, correlations between hemoglobin levels and FACIT-Fatigue total score were weak to moderate, as each employs a different method of data collection, and questionnaire scores typically lag change in biomarkers

Known groups analysis of the FACIT-Fatigue total score

- Known groups were evaluated using PGIS scores at Baseline and Day 168, as well as participants experiencing increase in hemoglobin ≥2 g/dL in absence of red blood cell transfusion at Day 168
- Results for APPLY-PNH using collapsed PGIS categories (**Table 7**) showed that participants in the "No symptoms/Mild" group had significantly higher FACIT-Fatigue scores at both timepoints compared to participants in the "Moderate" and "Severe/Very severe" group. FACIT-Fatigue total score was also able to distinguish between groups based upon hemoglobin increase at Day 168 (**Table 8**)
- Known groups analysis results for APPOINT-PNH were consistent with those for APPLY-PNH

Table 1. Demographic descriptive information at Baseline

	APPLY-PNH Statistic or n (%)	APPOINT-PNH Statistic or n (%)
Age (years)		
N	95	40
Mean (SD)	51.4 (16.8)	42.1 (15.8)
Median	53.0	38.5
Min, Max	20.0, 84.0	18.0, 81.0
Sex		
Male	65 (68.4%)	23 (57.5%)
Female	30 (31.6%)	17 (42.5%)

Abbreviations: Max=Maximum; Min=Minimum; N=total number; SD=standard deviation

Table 2. Testing for essential unidimensionality

	Day 140	Day 168
PUC	67.2%	67.2%
PECV	98.2%	97.1%

Abbreviations: PECV=Percentage explained common variance; PUC=Percentage uncontaminated correlations

Table 3. FACIT-Fatigue internal consistency reliability

	Chronbach's α		McDonald's ω	
	APPLY-PNH	APPOINT-PNH	APPLY-PNH	APPOINT-PNH
Baseline	0.947	0.943	0.909	No convergence*
Day 42	0.953	0.918	0.914	0.920
Day 126	0.949	0.937	No convergence*	0.927
Day 140	0.958	0.920	0.920	0.908
Day 168	0.959	0.895	0.929	0.905

*McDonald's ω could not be estimated because the factor analysis model used to derive the estimate did not converge.

Table 4. FACIT-Fatigue test-retest reliability

Analysis population	N*	ICC†	95% CI	
			Lower	Upper
APPLY-PNH (N=95)	16	0.934	0.824	0.976
APPOINT-PNH (N=40)	19	0.891	0.742	0.957

Abbreviations: CI=confidence interval; ICC=intra-class coefficient; N=number
*Only subjects with non-missing study instrument scores at both administrations were included in the analysis.
†The ICC was computed using the single measurement, absolute agreement, two-way mixed effects model.

Table 5. FACIT-Fatigue total score concurrent validity

Supplementary assessment	APPLY-PNH Range of correlations	APPOINT-PNH Range of correlations
EORTC-QLQ-C30 Physical function score*	0.75 to 0.79	0.71 to 0.85
EORTC-QLQ-C30 Role function score*	0.75 to 0.83	0.74 to 0.82
EORTC-QLQ-C30 Fatigue score*	-0.82 to -0.89	-0.78 to -0.87
EORTC-QLQ-C30 Dyspnea score*	-0.57 to -0.67	-0.10 to -0.62
EORTC-QLQ-C30 Global health status score*	0.71 to 0.83	0.53 to 0.75
EQ-5D-5L Mobility score†	-0.56 to -0.67	-0.45 to -0.58
EQ-5D-5L Self-care score†	-0.23 to -0.41	-0.26 to -0.43
EQ-5D-5L Usual activities score†	-0.72 to -0.84	-0.44 to -0.74
EQ-5D-5L Pain/discomfort score†	-0.27 to -0.55	-0.24 to -0.38
EQ-5D-5L Visual analog scale score†	0.66 to 0.76	0.41 to 0.71
PGIS‡	-0.85 to -0.94	-0.67 to -0.84

*EORTC QLQ-C30 scores range from 0 to 100. Higher scores on the physical functioning, role function, and global health status/quality of life scores indicate higher functioning and quality of life. Higher scores on the fatigue and dyspnea domains indicate higher symptom experiences (i.e., greater severity). †The EQ-5D-5L domain items are five-response ordinal items where higher scores represent worse health states. The EQ-5D-5L VAS ranges from 0 to 100 where higher scores represent better health states.
‡The PGIS asks participants to rate their overall symptoms of fatigue during the past seven days. The PGIS is rated on an ordinal scale ranging from 0 to 4 with higher scores indicating greater symptom severity.

Table 6. FACIT-Fatigue total score responsiveness to change

Supplementary assessment	n		Correlation	
	APPLY-PNH	APPOINT-PNH	APPLY-PNH	APPOINT-PNH
From Baseline to Day 42				
Change in EORTC-QLQ-C30 Fatigue score*	93	39	**-0.74**	**-0.78**
Change in PGIS†	93	39	**-0.81**	**-0.74**
Change in hemoglobin level‡	70	38	**0.26**	**0.17**
From Baseline to Day 168				
Change in EORTC-QLQ-C30 Fatigue score*	90	37	**-0.76**	**-0.78**
Change in PGIS†	90	37	**-0.74**	**-0.83**
Change in hemoglobin level‡	71	36	**0.37**	**-0.28**

Abbreviations: CI=confidence interval; n=number of participants included in analysis population;
*EORTC QLQ-C30 Fatigue scores range from 0 to 100. Higher scores on the symptom domains indicate higher symptom experiences. Negative change scores are indicative of improvement.
†The PGIS asks participants to rate their overall symptoms of fatigue during the past seven days. The PGIS is rated on an ordinal scale ranging from 0 to 4 with higher scores indicating greater symptom severity. Negative change scores are indicative of improvement.
‡Hemoglobin level is a biomarker derived from a blood sample analysis. An increase in hemoglobin in the study sample is indicative of PNH patient improvement.

Table 7. FACIT-Fatigue total score known groups analysis by collapsed PGIS response

Visit	PGIS group	n	FACIT-Fatigue total score		Mean difference*
			Mean (SD)	Median (95% CI)	
Baseline (N=95)	Entire sample	95	33.6 (11.3)	34.0	
	No symptoms/mild	44	42.3 (6.9)	43.5 (40.0, 47.0)	
	Moderate	37	28.5 (6.0)	27.0 (25.0, 31.0)	**-13.8**
	Severe/Very severe	14	20.2 (11.9)	22.5 (11.0, 29.0)	**-8.3**
Day 168 (N=90)	Entire sample	90	39.2 (11.2)	42.0	
	No symptoms/mild	67	44.0 (6.9)	45.0 (43.0, 47.0)	
	Moderate	17	27.9 (7.4)	29.0 (26.0, 32.0)	**-16.1**
	Severe/Very severe	6	16.7 (8.9)	19.0 (0.0, 24.0)	**-11.2**

Abbreviations: CI=confidence interval; N=number; SD=standard deviation
*Median difference calculated as the mean of the group minus the mean of the adjacent previous group.

Table 8. FACIT-Fatigue total score known groups analysis by increase in hemoglobin

Visit	Hemoglobin group	n	FACIT-Fatigue total score		Mean difference*
			Mean (SD)	Median (95% CI)	
Day 168	Achieved ≥2 g/dL increase	50	43.7 (7.7)	45.5 (43.0, 48.0)	
	Did not achieve ≥2 g/dL increase	18	34.8 (14.3)	38.5 (26.0, 45.0)	**-7.0**

Abbreviations: CI=confidence interval; N=number; SD=standard deviation
*Median difference calculated as the mean of the group minus the mean of the adjacent previous group.

Disclosures

Jeffrey McDonald, Samantha Linton, Ethan Arenson, Roger Lamoureux, and Gavin Dickie are employees of Adelphi Values;, which was compensated by Novartis for the conduct and reporting of the analyses described here. At the time the research was conducted, Gilbert Ngerano was an employee of Adelphi Values.

References

1. Rodriguez A, Reise SP, Haviland MG. Evaluating bifactor models: Calculating and interpreting statistical indices. Psychol Methods. 2016;21(2):137-150.
2. Ferrando PJ, Lorenzo-Seva U, Navarro-Gonzalez D. unival: An FA-based R Package For Assessing Essential Unidimensionality Using External Validity Information. The R Journal. 2019;11(1):401-415.
3. Muthén B, Kaplan D, Hollis M. On structural equation modeling with data that are not missing completely at random. Psychometrika. 1987;52(3):431-462.
4. Reise SP. The Rediscovery of Bifactor Measurement Models. Multivariate Behavioral Research. 2012;47(5):667-696.
5. Cronbach LJ. Coefficient Alpha and the Internal Structure of Tests. Psychometrika. 1951;16(3):297-334.
6. McDonald RP. Test theory: A unified approach. Mahwah, NJ: Lawrence Erlbaum Associates; 1999.
7. Hinkle DE, Jurs SG, Wiersma W. Applied statistics for the behavioral sciences. 2nd ed. Boston: Houghton Mifflin; 2003.
8. Cohen J. Statistical Power Analysis for the Behavioral Sciences. 2nd ed. New York: Taylor Francis; 2013.