Reliability, validity, and clinically important differences (CIDs) on the NCCN/FACT Bladder Symptom Index (NFBISI-18) among individuals with locally advanced or metastatic urothelial cancer (UC)

J. D. Peipert,[1] J. Chang,[2] S. Li,[2] A. Di Pietro,[2] P. Cislo,[2] J. C. Cappelleri,[2] D. Cella[1]

[1]Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, Chicago, IL, USA; [2]Pfizer, New York, NY, USA

JAVELIN Bladder 100

BACKGROUND

- The NFBISI-18 is a measure of advanced bladder cancer–specific symptoms composed of a total scale and 3 subscales representing physical disease–related symptoms (DRS-P), emotional disease–related symptoms (DRS-E), treatment side effects (TSE), and function/well-being (F/WB)
- Previous research provided evidence for the reliability and content validity of the NFBISI-18 scales.[1] However, a full psychometric evaluation of the complete 18-item format of this instrument has not been conducted to date
- In addition, CIDs have not been estimated for the NFBISI-18 scales
- Due to the dearth of UC-focused PRO measures, additional psychometric evaluation of the NBFISI-18 scales will enhance our capacity to advance patient-focused drug development in this population

APPROACH

- This study used the primary data-cut (October 21, 2019) from the JAVELIN Bladder 100 trial (NCT02603432).[2,3] JAVELIN Bladder 100 is a phase 3, parallel arm trial comparing first-line maintenance treatment with avelumab + best supportive care (BSC) to BSC alone in patients with locally advanced or metastatic UC who have not progressed with first-line platinum-containing chemotherapy
- The results presented here are from a post-hoc analysis focusing on the NFBISI-18 measure. Of the 700 patients enrolled in the trial, 651 patients with complete baseline responses to the NFBISI-18 scales were analyzed. All reliability and validity analyses, with the exception of test-retest reliability, used baseline data from JAVELIN Bladder 100
- The TSE scale was excluded since no treatment side effects are expected at baseline
- Tests for reliability and validity were based on recommendations from an International Society of Quality of Life Research consensus statement[4]

STATISTICAL ANALYSIS

- Coefficient α was used to estimate internal consistency reliability. Test-retest reliability was estimated with the absolute agreement intraclass correlation coefficient (ICC) using one-way random effects analysis of variance for patients with stable scores on the EQ-5D utility index (UI) and visual analog scale (VAS) from cycles 2 to 3[5]
- Convergent validity was tested by estimating Spearman's ρ correlations between the NFBISI-18 scales and the EQ-5D-5L UI and VAS scales. EQ-5D UI values were obtained with a crosswalk from the 5L to 3L version, and then UK weights were applied
- Known-groups validity was tested by anchoring mean differences in NFBISI-18 scales to groups of age (<65, ≥65 years), Eastern Cooperative Oncology Group performance status (ECOG PS), and number of medical conditions/syndromes at baseline, which includes both acute and chronic conditions (1-9, ≥10)
- To estimate CIDs, we calculated differences in means between categories of the known group anchors for which Cohen's d was >0.2 (at least a small effect).[6] To provide context for the CIDs, we calculated distributional properties of the scales (1/3 SD, 1/2 SD, 1 standard error of measurement [SEM])

SCOPE

- **This study aims to expand the psychometric evidence in use of the National Comprehensive Cancer Network (NCCN)/Functional Assessment of Cancer Therapy (FACT) Bladder Symptom Index (NFBISI-18) as a valid patient-reported outcome (PRO) measure in clinical trials of patients with locally advanced or metastatic UC**

CONCLUSIONS

- **This analysis demonstrated that the NFBISI-18 is a reliable and valid instrument to measure symptoms in patients with locally advanced or metastatic UC**
- **The CID estimates from this study can help clinicians and researchers understand what difference in patient symptoms is clinically meaningful, as measured by the NFBISI-18, to inform clinical practice**
- **In addition, the CID estimates can be useful in the planning of future trials, particularly in terms of effect size and sample size determination**
- **The next step in this research program is longitudinal psychometric evaluation, including evaluation of responsiveness to change and estimating responder definitions for the NFBISI-18 scales**

Correspondence: **John Devin Peipert**, john.peipert@northwestern.edu

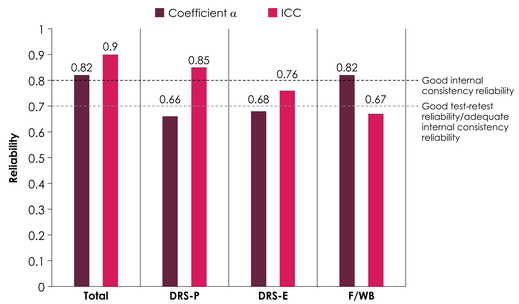
GET PLAIN LANGUAGE SUMMARY

Please scan this quick response (QR) code with your smartphone app or click here to view a plain language summary of the accepted scientific abstract

RESULTS

- Reliability estimates exceeded thresholds for reliability were generally considered acceptable for the majority of NFBISI-18 scales[7,8] (**Figure 1**)
- Spearman's correlations between the NFBISI-18 scales and the EQ-5D UI and VAS exceeded the threshold for large (>0.371)[6] for all but 1 pairwise correlation (DRS-E by EQ-5D VAS, ρ=0.32). These findings supported convergent validity for the NFBISI-18 scales. **Figure 2** shows each pairwise correlation
- **Table 1** shows results for the known-groups validity tests. Cohen's *d* for NFBISI-18 scale score differences between known groups ranged between 0.05 and 0.25 (age), 0.35 and 0.60 (ECOG PSR 0 vs 1), and 0.10 and 0.41 (number of comorbidities/symptoms). Most of these differences were statistically significant and supported known-group validity
- **Figure 3** shows estimates for CIDs, including lower and upper ranges. To provide context for these estimates, distributional values for each NFBISI-18 scale are included as well
- Recommendations for CIDs are as follows: total, 3-6 points; DRS-P, 2-3 points; DRS-E, 1 point; F/WB,1 point

Figure 1. Reliability estimates for NFBISI-18 scales

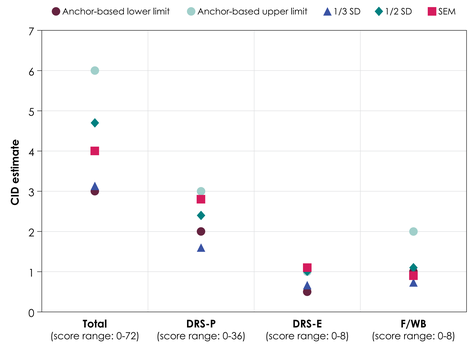


DRS-E, emotional disease-related symptoms; **DRS-P**, physical disease-related symptoms; **F/WB**, function/well-being; **ICC**, intraclass correlation coefficient; **NFBISI**, National Comprehensive Cancer Network/Functional Assessment of Cancer Therapy Bladder Symptom Index.

Figure 2. Convergent validity: Spearman's correlations between NFBISI-18 scales and EQ-5D

DRS-E, emotional disease-related symptoms; **DRS-P**, physical disease-related symptoms; **EQ-5D UI**, EQ-5D utility index; **EQ-5D VAS**, EQ-5D visual analogue scale; **F/WB**, function/well-being; **NFBISI**, National Comprehensive Cancer Network/Functional Assessment of Cancer Therapy Bladder Symptom Index.

Figure 3. Group-level CID estimates



CIDs are based on 2 anchor-based estimates (high and low); Distributional values given for context.
CID, clinically important differences; **DRS-E**, emotional disease-related symptoms; **DRS-P**, physical disease-related symptoms; **F/WB**, function/well-being; **SEM**, standard error of measurement.

Table 1. Known-groups validity

Anchor	N	Mean	Diff	p value	Effect size
NFBISI-18 total*					
Age, years					
<65	227	52.6	−0.7	0.41	−0.07
≥65	424	53.3			
ECOG PS					
0	397	55.3	5.6	<0.001	0.60
1	251	49.7			
No. of baseline health conditions or syndromes					
1-9	467	53.8	2.7	<0.001	0.29
≥10	184	51.1			
DRS-P†					
Age, years					
<65	227	27.3	0.2	0.78	0.04
≥65	424	27.1			
ECOG PS					
0	397	28.3	2.8	<0.001	0.58
1	251	25.5			
No. of baseline health conditions or syndromes					
1-9	467	27.5	1.2	0.003	0.25
≥10	184	26.3			
DRS-E‡					
Age, years					
<65	227	5.6	0.5	0.006	0.25
≥65	424	5.1			
ECOG PS					
0	397	5.7	0.7	<0.001	0.35
1	251	5.0			
No. of baseline health conditions or syndromes					
1-9	467	5.5	0.2	0.15	0.10
≥10	184	5.3			
F/WB§					
Age, years					
<65	227	4.5	−0.1	0.71	−0.05
≥65	424	4.6			
ECOG PS					
0	397	5.0	1.1	<0.001	0.50
1	251	3.9			
No. of baseline health conditions or syndromes					
1-9	467	4.7	0.4	0.04	0.18
≥10	184	4.3			

Diff, difference in adjacent group means; **DRS-E**, emotional disease-related symptoms; **DRS-P**, physical disease-related symptoms; **ECOG PS**, Eastern Cooperative Oncology Group performance status; **F/WB**, function/well-being; **NFBISI**, National Comprehensive Cancer Network/Functional Assessment of Cancer Therapy Bladder Symptom Index.
*Effect size calculated as Diff/pooled NFBISI-18 total SD (9.4).
†Effect size calculated as Diff/pooled DRS-P SD (4.8).
‡Effect size calculated as Diff/pooled DRS-E SD (2.0).
§Effect size calculated as Diff/pooled F/WB SD (2.2).

REFERENCES 1. Jensen SE, et al. J Support Oncol. 2013;11(2):86-93. **2.** Powles T, et al. N Engl J Med. 2020;383(13):1218-30. **3.** Powles T, et al. Ann Oncol. 2020;31(Suppl 4). Abstract 745P. **4.** Reeve BB, et al. Qual Life Res. 2013;22(8):1889-905. **5.** Qin S, et al. Qual Life Res. 2019;28(4):1029-33. **6.** Cohen J. Statistical Power Analysis for the Behavioral Sciences. Academic Press; 1988. **7.** Nunnally JC. Psychometric Theory. 2nd ed. McGraw-Hill; 1978. **8.** Fleiss JL, et al. Wiley series in probability and statistics. In: Statistical Methods for Rates and Proportions. John Wiley & Sons, Inc; 2004. **DISCLOSURES JDP** was supported by Pfizer for the current research and has been supported by AstraZeneca, FACIT Group, IMPAQ International, Janssen, and Veloxis for research unrelated to this project. **JC**, **SL**, **ADP**, **PC**, and **JCP** are employees of Pfizer. **DC** has been supported by Pfizer for the current research and has consulted with Merck & Co., Merck KGaA, Darmstadt, Germany, and Pfizer for work unrelated to this project. **ACKNOWLEDGMENTS** This study was sponsored by Pfizer as part of an alliance between Merck KGaA, Darmstadt, Germany, and Pfizer.

Abstract No. 408. Presented at the 2021 Genitourinary Cancers Symposium, February 11-13, 2021; Virtual.