Psychometric Validation of the EORTC QLQ-HCC18 in Patients with Previously Treated Unresectable Hepatocellular Carcinoma

Daniel Serrano[1], Lauren Podger[1], Gisoo Barnes[2], James Song[2], Boxiong Tang[2]

[1]Pharmerit (An Open Health Company), [2]BeiGene, Ltd.

Virtual ISPOR 2021
17-20 May 2021, Virtual Conference

BACKGROUND

- Hepatocellular carcinoma (HCC) accounts for 85-90% of all reported cases of liver cancer and is the 4th most common cause of cancer-related death worldwide (1).
- The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire Hepatocellular Carcinoma 18-question module (EORTC QLQ-HCC18) was developed specifically to assess symptom burden and impact on HRQoL in people with HCC (2-4).
  - However, there are limited published data on the psychometric properties of the QLQ-HCC18 within a previously untreated unresectable HCC population.
- NCT03419897 is an open-label, international, Phase 2 clinical trial assessing efficacy and safety of the monoclonal antibody BGB-A317 in adult patients with previously treated unresectable 2nd/3rd line HCC.
- The objective of this study was to evaluate the psychometric properties of the QLQ-HCC18 within a previously treated, unresectable HCC clinical trial population that was distinct from the population on which the measure was developed.

METHODS

- Analyses were conducted using data from the NCT03419897 trial.
  - Enrolled patients received 200mg of tislelizumab (BGB-A317) intravenously every 3 weeks for a total of ≥3 treatment cycles.
  - The EORTC Quality of Life Questionnaire Core 30 (EORTC QLQ-C30) and QLQ-HCC18 instruments were assessed at baseline, week 3 and week 9 follow-up visits. At each treatment cycle visit, administration of the HRQoL instruments occurred prior to any clinical activities or dosing.
- Per US FDA guidance (5,6), psychometric analysis of the QLQ-HCC18 included:
  - Internal consistency reliability evaluates score reliability by assessing the strength with which each item measures an assumed single domain and was assessed for the multi-item QLQ-HCC18 scales at baseline using Cronbach's alpha.
  - Test-retest reliability measures the degree to which an instrument is capable of reproducing scores across time in subjects whose condition has not changed and was assessed for the QLQ-HCC18 scores between baseline and week 3 using the two-way random intraclass correlation coefficients (ICC) anchored to no change on the QLQ-C30 Global Health/QoL (GHS) domain.
  - Concurrent validity was assessed between the QLQ-HCC18 and QLQ-C30 scores (the latter consisting of 16 scales and single items) at baseline using Spearman correlations.
  - Known-groups validity was assessed for the QLQ-HCC18 scores at baseline using analysis of variance (ANOVA) stratified on geographic region, line of therapy, and viral hepatitis status.
- Ability to detect change was assessed for the QLQ-HCC18 change scores between baseline and week 9 using analysis of covariance (ANCOVA) stratified on the QLQ-C30 GHS anchor groups (operationalized as improvement [>0-point change from baseline to week 9], maintenance [no change], or deterioration [<0-point change]), controlling for age, gender, region, and baseline QLQ-HCC18 mean change.
- Meaningful within-patient change (MWPC) was assessed for the QLQ-HCC18 change scores between baseline and week 9 stratified on the QLQ-C30 GHS anchor-based thresholds described above. Estimates of mean change were validated by visualizing differences in cumulative proportions achieving the point estimates via empirical cumulative distribution functions (eCDFs).

RESULTS

- A total of 249 patients were enrolled in BGB-A317-208 trial and received at least 1 dose of tislelizumab (BGB-A317); see Table 1.
- A single patient who did not contribute QLQ-HCC18 data at baseline was excluded, leaving a final sample of 248 patients for the psychometric analysis.

Table 1. Patient demographics and clinical characteristics

Characteristic	Total Sample (N = 249)	Line of Therapy	
		Second-line (n = 138)	Third-line + (n = 111)
Age (years)			
Mean (SD)	60.3 (12.5)	60.2 (13.7)	60.4 (10.9)
Gender, n (%)			
Male	217 (87.1)	121 (87.7)	96 (86.5)
Female	32 (12.9)	17 (12.3)	15 (13.5)
ECOG performance status at baseline, n (%)			
0	129 (51.8)	70 (50.7)	59 (53.2)
1	120 (48.2)	68 (49.3)	52 (46.8)
Child-Pugh classification at baseline, n (%)			
A	248 (99.6)	138 (100.0)	110 (99.1)
B	1 (0.4)	0 (0.0)	1 (0.9)
Hepatitis virus infection, n (%)			
Uninfected	90 (36.1)	46 (33.3)	44 (39.6)
Hepatitis B only	123 (49.4)	71 (51.4)	52 (46.8)
Hepatitis C only	31 (12.4)	20 (14.5)	11 (9.9)
Coinfected	5 (2.0)	1 (0.7)	4 (3.6)

SD: standard deviation; ECOG: Eastern Cooperative Oncology Group.

Reliability

- Internal consistency reliability: the QLQ-HCC18 fatigue, nutrition, and index domains demonstrated acceptable internal consistency at baseline based on the pre-specified threshold of α ≥0.70; see Table 2.

Table 2. Internal consistency of the QLQ-HCC18 at baseline

QLQ-HCC18 Domain	Cronbach's Alpha (N = 248)
Fatigue (3 items)	**0.71**
Body image (2 items)	0.53
Jaundice (2 items)	-0.07
Nutrition (5 items)	**0.75**
Pain (2 items)	0.45
Fever (2 items)	0.23
Index (18 items)	**0.88**

Note: Bold text indicates those estimates that reached the acceptable threshold (≥0.70).

Validity

- Concurrent validity: the QLQ-HCC18 fatigue domain achieved the pre-specified criterion of r ≥|0.4| defining acceptable concurrent validity for 13 out of 16 (81%) correlations, whereas the index domain achieved the pre-specified criterion for 15 out of 16 (94%) correlations.
- Known-groups validity:
- For the QLQ-HCC18 fatigue, body image, jaundice, and index domains, patients in Europe reported significantly higher mean scores (i.e., worse symptoms or reduced HRQoL) compared with patients in Asia.
- For the QLQ-HCC18 body image domain, patients in the viral hepatitis negative group reported a significantly higher mean score compared with those patients in the viral hepatitis positive group.
- For the QLQ-HCC18 jaundice domain, patients in the 3rd-line or greater therapy group reported a significantly higher mean score compared with those patients in the 2nd-line therapy group.

Responsiveness

- Ability to detect change: clear differentiation of the QLQ-HCC18 change scores between improvement and maintenance anchor groups were observed for body image, fatigue, pain, and index domains, whereas differentiation between deterioration and maintenance anchor groups were observed for fever and fatigue domains.
- MWPC: point estimates defining improvement for the QLQ-HCC18 fatigue and index domains were -7.18 and -4.07, respectively; point estimates defining deterioration were 5.34 and 3.16, respectively. See Figures 1 and 2.

Figure 1. Empirical cumulative distribution function of the QLQ-HCC18 fatigue domain change score from baseline to week 9 stratified on the QLQ-C30 GHS anchor groups

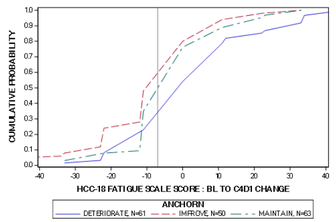
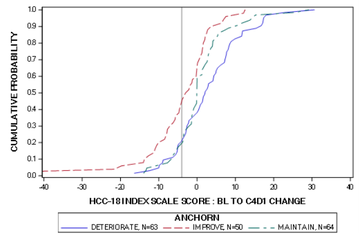


Figure 2. Empirical cumulative distribution function of the QLQ-HCC18 index domain change score from baseline to week 9 stratified on the QLQ-C30 GHS anchor groups



DISCUSSION

- The EORTC QLQ-HCC18 fatigue and index domains consistently demonstrated robust psychometric properties, supporting the use of these domains as suitable patient-reported endpoints within a previously treated, unresectable HCC patient population.
  - However, the remaining QLQ-HCC18 domains did not consistently demonstrate optimal measurement properties in this population.
- The ability to detect change and MWPC analyses demonstrated that a detectable improvement was observed in this trial and the QLQ-HCC18 fatigue domain scores sensitively detected the effect of tislelizumab.
- These results are consistent with previous QLQ-HCC18 validation studies (2,7).

LIMITATION

- The GHS anchor employed in MWPC was the only anchor collected in the trial, ideally a global symptom impression of severity would be employed in MWPC.

REFERENCES

1. International Agency for Research on Cancer, World Health Organization. Cancer today (https://gco .iarc .fr/ today/ home).
2. Chie WC, Blazeby JM, Hsiao CF, Chiu HC, Poon RT, Mikoshiba N, et al. International cross-cultural field validation of an European Organization for Research and Treatment of Cancer questionnaire module for patients with primary liver cancer, the European Organization for Research and Treatment of Cancer quality-of-life questionnaire HCC18. *Hepatology*. 2012;55(4):1122-9.
3. Mikoshiba N, Tateishi R, Tanaka M, Sakai T, Blazeby JM, Kokudo N, et al. Validation of the Japanese version of the EORTC hepatocellular carcinoma-specific quality of life questionnaire module (QLQ-HCC18). *Health and quality of life outcomes*. 2012;10:58.
4. Yang Z, Wan C, Li W, Cun Y, Meng Q, Ding Y, et al. Development and Validation of the Simplified Chinese Version of EORTC QLQ-HCC18 for Patients with Hepatocellular Carcinoma. *Cancer investigation*. 2015;33(8):340-6.
5. Food and Drug Administration Guidance for Industry. Patient-reported outcome measures: use in medical product development to support labeling claims. Silver Spring; 2009. https://www.fda.gov/downloads/Drugs/.../Guidances/UCM193282.pdf.
6. Food and Drug Administration (FDA), 2018. Methods to Identify What is Important to Patients & Select, Develop or Modify Fit-for-Purpose Clinical Outcomes Assessments. https://www.fda.gov/media/116277/download.
7. Li L, Mo FK, Chan SL, Hui EP, Tang NS, Koh J, et al. Prognostic values of EORTC QLQ-C30 and QLQ-HCC18 index-scores in patients with hepatocellular carcinoma - clinical application of health-related quality-of-life data. *BMC cancer*. 2017;17(1):8.

Acknowledgments

The authors wish to thank Jason Allaire, PhD of Generativity Solutions Group for his assistance in medical writing. This assistance was funded by BeiGene, Ltd.

Declaration of Funding

The NCT03419897 trial as well as this analysis was funded by BeiGene, Ltd.