Psychometric Validation of the EORTC QLQ-OES18 in Second-Line Esophageal Squamous Cell Cancer Patients Treated With Tislelizumab vs Chemotherapy

PCR243

Lauren Podger,[1] Daniel Serrano,[2] Liyun Li,[3] Lin Zhan,[3] Boxiong Tang,[3] Gisoo Barnes[3]

[1]Open Health Group, London, UK. [2]The Psychometrics Team, WY, USA. [3]BeiGene, Ltd, Emeryville, CA, USA

BACKGROUND

- Esophageal squamous cell carcinoma (ESCC), the most common histological subtype of esophageal cancer, frequently results in a high burden for patients at diagnosis, as well as reductions in health-related quality of life (HRQOL) due to esophageal obstruction throughout the disease course, including during treatment[1-3]
- A better understanding of ESCC-specific symptoms, functioning, and HRQOL is needed for patients with advanced or metastatic ESCC
  - This need is especially relevant given the current paradigm shift in ESCC treatment options away from second-line chemotherapy as standard of care[4-6]
- The European Organization for Research and Treatment of Cancer Quality of Life Questionnaire – Oesophageal Cancer 18-question module (EORTC QLQ-OES18) has previously demonstrated clinical validity and is frequently used[7]; however, there are limited published psychometric data for patients with advanced or metastatic ESCC

OBJECTIVE

- To evaluate the measurement properties of the EORTC QLQ-OES18 instrument in a clinical trial population to establish evidence that it is fit for purpose in patients with advanced or metastatic ESCC

METHODS

Study Design

- Analyses were conducted using data from RATIONALE-302 (NCT03430843), a global, open-label, randomized, phase 3 study that investigated the efficacy and safety of tislelizumab, an anti–PD-1 monoclonal antibody, vs investigator-chosen chemotherapy as second-line treatment for patients aged ≥18 years with advanced or metastatic ESCC whose disease had progressed after first-line systemic therapy[8]
  - Compared to ICC, tislelizumab was found to prolong OS (median of 8.6 versus 6.3 months; hazard ratio [HR] 0.70, 95% confidence interval [CI] 0.57–0.85, P=0.0001) and was associated with a higher objective response rate (ORR; 20.3% versus 9.8%)
  - Tislelizumab was also found to have a more durable anti-tumor response compared to ICC (median of 7.1 versus 4.0 months)

Study Measures

- EORTC QLQ-OES18 and EORTC Quality of Life Questionnaire – Core 30 (QLQ-C30)[9] scores were assessed at baseline and at the week 3 follow-up visit
- The 18-item QLQ-OES18 is designed to assess HRQOL in patients with esophageal cancer
  - It consists of 4 multi-item symptom scales (dysphagia, eating, reflux, and pain), 6 symptom single items (swallowing saliva, choking when swallowing, dry mouth, trouble with taste, trouble with coughing, and trouble with talking), and a composite symptom index scale
  - Transformed scores range from 0 to 100; higher scores indicate worse symptoms or reduced HRQOL
- The 30-item QLQ-C30 is a generic measure for evaluating HRQOL across a range of issues in patients with cancer
  - It includes a total of 15 domains (global health status/quality of life [GHS/QOL], physical functioning, role functioning, emotional functioning, cognitive functioning, and social functioning) and symptoms (fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties)?
  - Transformed scores range from 0 to 100; higher scores on the GHS/QOL and functional scales represent higher global QOL or functioning level, whereas higher symptom scores represent worse symptoms or problems

Statistical Analyses

- Psychometric validation of the QLQ-OES18 included tests of reliability (internal consistency and test-retest reliability) and construct validity (convergent/discriminant and known-groups validity) (**Table 1**)

Table 1. Summary of Psychometric Analyses of QLQ-OES18

Property	Analysis period	Definition	Test	Success criterion
Interitem correlations	Baseline	Polychoric correlation	No test; point estimate reported	|r| ≥0.40
Internal consistency	Baseline	Cronbach α	No test; point estimate reported	α ≥0.70
Test-retest reliability	Baseline to week 3	ICC(A,1)	No test; point estimate reported	ICC ≥0.70
Concurrent validity	Baseline	Spearman correlation	No test; point estimate reported	|r| ≥0.40
Known-groups validity	Baseline	Mean difference; 95% CI, *P* value, and ω^2 effect size	ANOVA	*P* <.05; effect size >5%
Meaningful within-patient change	Baseline	Threshold: median anchor-based change score		
eCDF plotted	Separation of anchor group eCDFs at the threshold location	Separation of eCDFs		

Analyses were conducted using transformed scores on both the QLQ-OES18 and QLQ-C30.
ANOVA, analysis of variance; eCDF, empirical cumulative distribution function; ICC, intraclass correlation coefficient; QLQ-OES18, Quality of Life Questionnaire – Oesophageal Cancer 18-question module.

RESULTS

- Overall, 512 patients were randomized to either tislelizumab or chemotherapy
- The cohort had an average age of 61.5 years and was mostly male (84.4%), Asian (79.7%), and non-Hispanic (98.4%); all baseline characteristics were balanced across treatment groups
  - A total of 23 patients were missing patient-reported outcome data at baseline and were excluded from the current psychometric analyses
- Response patterns for the QLQ-OES18 at baseline showed that all 18 items had floor effects, whereby the "not at all" category was most frequently selected. However, this is reflective of the low severity of disease in the patient sample at baseline and was not unexpected. In addition, 98% of unique interitem correlations were ≤0.8, suggesting that the items independently contributed data

Reliability

- Three of the 4 QLQ-OES18 domains demonstrated acceptable internal consistency (α = 0.87, 0.77, and 0.71 for dysphagia, eating, and pain, respectively); the index score also demonstrated acceptable internal consistency (α = 0.78)
  - Only the reflux domain did not meet the prespecified success criterion, although it was within rounding distance (α = 0.67)
- For test-retest reliability, the ICC(A,1) estimates ranged between 0.41 and 0.78, exceeding the prespecified success criterion for the dysphagia, eating, and pain domains (ICC = 0.78, 0.77, and 0.76, respectively) and the index score (ICC = 0.78)

Validity

- Convergent validators based on the QLQ-C30 scores were expected to correlate with QLQ-OES18 scores with an absolute value of ≥0.40, whereas discriminant validators were not expected to correlate with the QLQ-OES18
- Associations between the QLQ-OES18 domain scores and the QLQ-C30 validators were generally as expected (**Table 2**)
  - For example, a strong correlation was observed between the QLQ-OES18 pain score and QLQ-C30 pain score; a moderate correlation was observed between the QLQ-OES18 trouble-with-taste score and QLQ-C30 appetite score
  - None of the QLQ-OES18 domain scores correlated with the QLQ-C30 discriminant validators (QLQ-C30 insomnia, diarrhea, or financial difficulties scores)

Table 2. QLQ-OES18 Concurrent Validity at Baseline

	QLQ-OES18 domain										
QLQ-C30 validator	Dry mouth	Eating	Trouble with coughing	Dysphagia	Pain	Reflux	Swallowing saliva	Choke when swallowing	Trouble with taste	Trouble with talking	Index scale
Physical functioning	−0.32	**−0.46**	−0.36	−0.17	**−0.48**	−0.26	−0.30	−0.16	−0.36	−0.23	**−0.52**
Role functioning	−0.30	**−0.45**	−0.27	−0.13	**−0.43**	−0.29	−0.21	−0.13	−0.33	−0.22	**−0.46**
Emotional functioning	−0.29	**−0.45**	−0.32	−0.14	**−0.45**	−0.31	−0.25	−0.23	−0.32	−0.19	**−0.49**
Cognitive functioning	−0.34	**−0.42**	−0.34	−0.19	**−0.49**	−0.34	−0.32	−0.18	−0.39	−0.28	**−0.55**
Social functioning	−0.28	−0.36	−0.27	−0.10	−0.31	−0.29	−0.23	−0.15	−0.30	−0.20	**−0.42**
Fatigue	0.38	**0.51**	0.39	0.18	**0.55**	0.39	0.27	0.21	0.37	0.27	**0.59**
Nausea and vomiting	0.16	**0.49**	0.38	0.17	**0.46**	**0.46**	0.14	0.23	0.26	0.18	**0.49**
Pain	0.30	0.35	0.30	0.08	**0.59**	0.33	0.15	0.17	0.36	0.20	**0.46**
Dyspnea	0.29	0.32	**0.42**	0.12	0.33	0.26	0.23	0.14	0.21	0.28	**0.44**
Insomnia	0.19	0.30	0.27	0.08	0.35	0.32	0.18	0.13	0.24	0.17	0.37
Appetite loss	0.30	**0.55**	0.24	0.18	**0.42**	0.38	0.16	0.15	**0.48**	0.14	**0.51**
Constipation	0.29	0.31	0.30	0.10	0.34	0.34	0.18	0.11	0.34	0.23	**0.42**
Diarrhea	0.10	0.18	0.16	0.10	0.18	0.26	0.03	0.06	0.04	0.06	0.20
Financial difficulties	0.19	0.10	0.10	−0.05	0.10	0.21	0.08	0.19	0.14	0.12	0.19
GHS/QOL	−0.28	**−0.46**	−0.30	−0.20	**−0.45**	−0.28	−0.30	−0.18	−0.33	−0.20	**−0.51**
Index scale	**0.41**	**0.58**	**0.46**	0.17	**0.61**	**0.50**	0.29	0.25	**0.47**	0.31	**0.68**

Bold text indicates estimates that reached the prespecified threshold for acceptable correlations (|r| ≥0.40).
GHS/QOL, global health status/quality of life scale; QLQ-C30, Quality of Life Questionnaire – Core 30; QLQ-OES18, Quality of Life Questionnaire – Oesophageal Cancer 18-question module.

CONCLUSIONS

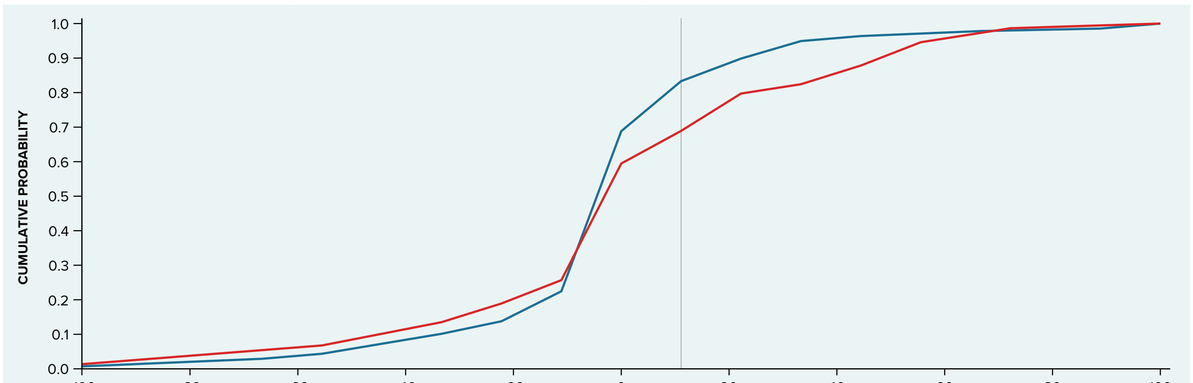
- Many patients enrolled in RATIONALE-302 reported only minimal symptoms at baseline, which resulted in restriction of range across several QLQ-OES18 domain and item scores
- Overall, the collection of psychometric and statistical evidence indicated that the QLQ-OES18 was able to reliably and validly measure symptom severity in the RATIONALE-302 population
- Specifically, the dysphagia domain consistently demonstrated robust psychometric properties, supporting its use as a patient-reported endpoint in trials evaluating the efficacy of novel treatments in people with advanced or metastatic ESCC

- The mean differences in QLQ-OES18 scores were calculated between the known groups defined by geographic region (Asia vs US/EU), Eastern Cooperative Oncology Group (ECOG) performance status (0 vs 1), and QLQ-C30 GHS/QOL scale scores (ratings of 1-6 vs 7; items 29 and 30 were treated as separate validators)
  - It was hypothesized that US/EU patients would report worse symptoms vs Asian patients, patients with an ECOG performance status of 1 would report worse symptoms vs those with a performance status of 0, and patients who selected "excellent" on the QLQ-C30 GHS/QOL scale items would report lower symptom severity vs those who selected any other response option
- A total of 40 of 44 (90.9%) comparisons demonstrated the hypothesized direction of effect, suggesting that the expected differences in the QLQ-OES18 domain scores between prespecified groups were observed (**Table 3**)

Meaningful Within-Patient Change

- Anchor-based MWPC thresholds were estimated
- The anchor used was the QLQ-C30 GH scale categorized into ≥2-point deterioration/improvement, 1-point deterioration/improvement, and no change
- For the OES-18 dysphagia domain (**Figure 1**), the MWPC threshold was based on the median change score for the 2+ point deterioration anchor group:
  - 11.11 points
- Treatment arm–stratified eCDFs were separated at the threshold location
- The risk of deterioration was approximately 15% lower in the tislelizumab group using this threshold

Figure 1. eCDF of OES-18 Dysphagia Domain From Baseline to C4D1 Change



BL, baseline; C4D1, cycle 4 day 1; eCDF, empirical cumulative distribution functions; ICC, intraclass correlation coefficient; QLQ-OES18, Quality of Life Questionnaire – Oesophageal Cancer 18-question module.

Table 3. QLQ-OES18 Known-Groups Validity at Baseline by QLQ-C30 Validators

QLQ-OES18 domain	Contrast	Group mean difference	95% CI	*P* value	Effect size (ω^2)
Dry mouth	Region	2.09	−2.99 to 7.16	.4196	0.394
	ECOG	7.91	3.29-12.53	.0008	0.406
	GHS/QOL item 29	−11.82	−18.95 to −4.68	.0012	0.405
	GHS/QOL item 30	−8.68	−14.6 to −2.76	.0041	0.402
Eating	Region	11.30	6.77-15.84	<.0001	0.477
	ECOG	6.10	1.88-10.33	.0048	0.459
	GHS/QOL item 29	−12.77	−19.25 to −6.29	.0001	0.466
	GHS/QOL item 30	−11.34	−16.69 to −6.00	<.0001	0.469
Trouble with coughing	Region	1.22	−3.65 to 6.08	.6236	0.238
	ECOG	4.43	−0.03 to 8.89	.0515	0.243
	GHS/QOL item 29	−8.79	−15.65 to −1.93	.0122	0.247
	GHS/QOL item 30	−8.45	−14.11 to −2.78	.0036	0.251
Dysphagia	Region	0.58	−7.26 to 8.42	.8845	0.502
	ECOG	−0.87	−8.12 to 6.38	.8137	0.501
	GHS/QOL item 29	0.17	−11.02 to 11.36	.9763	0.501
	GHS/QOL item 30	−6.05	−15.29 to 3.19	.1988	0.503
Pain	Region	5.60	1.92-9.29	.0030	0.349
	ECOG	4.31	0.90-7.73	.0134	0.345
	GHS/QOL item 29	−11.15	−16.36 to −5.94	<.0001	0.360
	GHS/QOL item 30	−10.76	−15.04 to −6.48	<.0001	0.368
Reflux	Region	−0.26	−4.25 to 3.72	.8974	0.329
	ECOG	5.64	1.99-9.29	.0025	0.341
	GHS/QOL item 29	−6.51	−12.16 to −0.85	.0242	0.336
	GHS/QOL item 30	−6.11	−10.78 to −1.44	.0104	0.338
Swallowing saliva	Region	7.16	1.87-12.45	.0081	0.197
	ECOG	8.34	3.47-13.21	.0008	0.204
	GHS/QOL item 29	−2.69	−10.29 to 4.91	.4870	0.186
	GHS/QOL item 30	−8.24	−14.48 to −1.99	.0098	0.196
Choke when swallowing	Region	−2.43	−7.05 to 2.19	.3021	0.322
	ECOG	4.21	−0.05 to 8.47	.0528	0.326
	GHS/QOL item 29	−8.41	−14.97 to −1.85	.0121	0.329
	GHS/QOL item 30	−4.77	−10.21 to 0.67	.0857	0.324
Trouble with taste	Region	7.60	2.67-12.52	.0026	0.214
	ECOG	4.14	−0.41 to 8.68	.0743	0.204
	GHS/QOL item 29	−10.28	−17.26 to −3.31	.0039	0.213
	GHS/QOL item 30	−10.86	−16.6 to −5.12	.0002	0.221
Trouble with talking	Region	0.94	−3.83 to 5.71	.6992	0.193
	ECOG	7.45	3.09-11.81	.0008	0.210
	GHS/QOL item 29	−6.97	−13.75 to −0.20	.0438	0.199
	GHS/QOL item 30	−2.89	−8.51 to 2.74	.3137	0.194
Index scale	Region	3.26	0.42-6.10	.0246	0.608
	ECOG	5.13	2.56-7.70	.0001	0.615
	GHS/QOL item 29	−7.89	−11.86 to −3.92	.0001	0.615
	GHS/QOL item 30	−7.78	−11.04 to −4.52	<.0001	0.620

The geographic region validator compared patients in Asia vs those in the US/EU. The ECOG performance status validator compared a status of 0 vs 1. The QLQ-C30 GHS/QOL item 29 validator compared ratings of 1 to 6 vs 7. The QLQ-C30 GHS/QOL 30-item validator compared ratings of 1 to 6 vs 7.
ECOG, Eastern Cooperative Oncology Group; GHS/QOL, global health status/quality of life scale; QLQ-C30, Quality of Life Questionnaire – Core 30; QLQ-OES18, Quality of Life Questionnaire – Oesophageal Cancer 18-question module.

REFERENCES

1. Sunde B, et al. *BMC Cancer.* 2021;21(1):1277.
2. Davis LE, et al. *Dis Esophagus.* 2020;33(8):doz108.
3. Wang QL, et al. *Clin Epidemiol.* 2018;10:717-728.
4. Huang J, et al. *Lancet Oncol.* 2020;21(6):832-842.
5. Kato K, et al. *Lancet Oncol.* 2019;20(11):1506-1517.
6. Kojima T, et al. *J Clin Oncol.* 2020;38(35):4138-4148.
7. Blazeby JM, et al. *Eur J Cancer.* 2003;39(10):1384-1394.
8. Shen L, et al. *J Clin Oncol.* 2022;40(26):3065-3076.
9. Aaronson NK, et al. *J Natl Cancer Inst.* 1993;85(5):365-376.

DISCLOSURES

LP is an employee of Open Health Group, and DS is an employee of The Psychometrics Team; both were paid for their services as consultants for BeiGene, Ltd. TX, LL, LZ, BT, and GB are employees of BeiGene, Ltd.

ACKNOWLEDGMENTS

The authors wish to thank Tianyu Xia for their contributions to the poster. Medical writing and editorial assistance were provided by Jason Allaire, PhD, of Generativity Solutions Group and funded by BeiGene, Ltd.

CORRESPONDENCE: Leyla Mohseninejad, leyla.mohseninejad@beigene.com

Presented at ISPOR Europe 2023; November 12-15, 2023; Copenhagen, Denmark