

European Organization for Research and Treatment of Cancer Quality of Life Questionnaire–Head and Neck Module, updated version: Preliminary psychometric data from Serbian laryngectomized patients

Sanja Krejovic Trivic, MD, MSc,¹ Aleksandar Trivic, MD, PhD,¹ Susanne Singer, PhD,² Jovica Milovanovic, MD, PhD,¹ Predrag Stankovic, MD, PhD,¹ Anton Mikic, MD, PhD,¹ Milan Vukasinovic, MD, MSc,¹ Vladimir Djordjevic, MD, PhD,¹ Ana Jotic, MD, MSc,¹ Miljan Folic, MD,¹ Dejan Stevanovic, MD, PhD^{3*}

¹Clinic of Otorhinolaryngology and Maxillofacial Surgery, Faculty of Medicine, University of Belgrade, Belgrade, Serbia, ²Institute of Medical Biostatistics, Epidemiology and Informatics (IMBEI), University Medical Centre Mainz, Mainz, Germany, ³Clinic for Neurology and Psychiatry for Children and Youth, Belgrade, Serbia.

Accepted 13 May 2015

Published online 18 July 2015 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/hed.24127

ABSTRACT: *Background.* We provided preliminary psychometric data for the European Organization for Research and Treatment of Cancer (EORTC) Quality of Life Questionnaire–Head and Neck Module, updated version (QLQ-H&N43) from a group of Serbian laryngectomized patients.

Methods. The study included 170 subjects. The QLQ-H&N43 is a 43-item questionnaire, with 12 multi-item scales and 7 single-item symptom scales. All subjects also completed the Quality of Life Questionnaire-Core 30-questions (QLQ-C30).

Results. Good internal consistency (Cronbach's α of above 0.7) was found for 5 of the 7 scales. All QLQ-H&N43 scales correlated negatively as predicted with all QLQ-C30 functioning scales. The correlations with the QLQ-C30 symptoms supported discriminant validity, with only one exception: the head and neck social eating scale overlapped with the

QLQ-C30 pain scale. For 14 of 19 QLQ-H&N43 scale scores, significant known-group differences were observed between those who differ in type of laryngectomy, adjuvant therapy, or 5-year survival.

Conclusion. Preliminary evidence suggests that a great majority of the QLQ-H&N43 scales have acceptable internal consistency and promising construct validity, but more research studies are needed with other cancer groups to extend these findings. © 2015 Wiley Periodicals, Inc. *Head Neck* 38: E916–E924, 2016

KEY WORDS: psychometrics, laryngeal cancer, laryngectomy, European Organisation for Research and Treatment of Cancer (EORTC), quality of life

INTRODUCTION

The concept of health-related quality of life (QOL) has emerged in the era when longer survival and improved living for people with chronic conditions are the paramount goals of the health care system. As a multidimensional concept simultaneously assessing physical, mental, and social functioning domains, QOL goes beyond other measures of population health, such as life expectancy and causes of death, and it focuses on the impacts of patients' current health status on their everyday well-being and functioning.¹ Over the past 2 decades, research on survival and life improvement in people with cancer has favored the QOL concept as a complementary outcome measure.²

Taking a systematic approach to QOL assessments, the European Organization for Research and Treatment of Cancer (EORTC) developed a generic questionnaire – the EORTC Quality of Life Questionnaire-Core 30-questions (QLQ-C30) for QOL assessments in all patients with

cancer, as well as various QOL assessment modules that cover the specific symptoms, treatment side-effects, and functional problems across different cancer types.^{3,4} One of these modules is the Quality of Life Questionnaire–Head and Neck Module 35-questions (QLQ-H&N35) developed for patients with head and neck cancers.^{5–7} A recent review showed that the QLQ-H&N35 frequently demonstrated good acceptance and feasibility when used in clinical trials and observational studies, and that it can also be successfully implemented in clinical practice.⁸ Moreover, the use of the QLQ-H&N35 in 19 languages across 26 countries indicates broad cross-cultural acceptance.⁸

Nevertheless, this recent review highlighted some of the module's psychometric problems, such as high percentage of missing values and low internal consistency of some scales. This initiated a joint taskforce by the EORTC Head and Neck Cancer Group and the EORTC Quality of Life Group to revise the module.^{8,9} Considering that the module was first developed in 1994, the group first conducted a literature review to identify potentially relevant new issues for patients with head and neck cancers.⁹ Afterward, a provisional updated module was created using the EORTC item bank. This provisional module contained 60 items, which were pilot-tested with

*Corresponding author: D. Stevanovic, Department of Psychiatry, Clinic for Neurology and Psychiatry for Children and Youth, Dr. Subotic 6a 11000, Belgrade, Serbia. E-mail: dejanstevanovic@eunet.rs

respect to their ease of understanding, comprehensiveness, and applicability on a group of patients with head and neck cancer.¹⁰ As a result, it was suggested that the QLQ-H&N updated module should contain 43 items, with 27 old and 16 new or revised items.⁹ Twelve 2-to-5 item symptom scales were proposed: QLQ-H&N pain (4 items), QLQ-H&N swallowing (4 items), QLQ-H&N speech problems (5 items), QLQ-H&N senses problems (2 items), QLQ-H&N social eating (4 items), QLQ-H&N problems with teeth (3 items), QLQ-H&N body image (3 items), QLQ-H&N skin problems (3 items), QLQ-H&N sexuality (2 items), QLQ-H&N dry mouth/sticky saliva (2 items), QLQ-H&N shoulder problems (2 items), and QLQ-H&N anxiety (2 items). Seven single-item symptom scales were proposed: QLQ-H&N social contact, QLQ-H&N opening mouth, QLQ-H&N coughing, QLQ-H&N lymphedema, QLQ-H&N weight loss, QLQ-H&N problems with wound healing, and QLQ-H&N neurological problems. All items have a Likert-like response format (not at all = 1; a little = 2; quite a bit = 3; and very much = 4). Scores for all scales and single items were calculated by linear transformation of raw scores into a 0 to 100 score, with 100 representing heavy burden. The EORTC Quality of Life Group and the Head and Neck Cancer Group have plans to conduct a larger field study to test its psychometric properties and confirm or revise the proposed scale structure.

In order to provide first insights into its measurement properties when used in another patient sample in another country, the purpose of the present study was to provide a preliminary evaluation of psychometric properties for the QLQ-H&N43 from a group of Serbian patients who underwent laryngectomy for laryngeal carcinoma.

MATERIALS AND METHODS

Participants and procedures

The QLQ-H&N43 was translated by the authors from English into Serbian following a standardized forward-backward procedure developed by the EORTC¹¹ in collaboration with the EORTC group. After native Serbian speakers with a clinical background had checked the translation, the Serbian version was pretested in a group of 5 laryngectomized patients, who were native Serbian. Initially, all patients completed the questionnaire independently. Finally, they participated in an independent semistructured interview. During the interviews, the following aspects of the translation were explored: precision (ie, an item should measure only 1 concept without overlapping with the concept of any other items), clarity (ie, an item should be formulated clearly), effectiveness (ie, an item should evoke a desired result, so the subject may easily select only 1 response option), relevance (ie, an item should be relevant as a QOL item for this group), and appropriateness (ie, an item should be not offensive or provocative). Except for necessary linguistic validation, no translated items were subsequently changed, added, or removed. Once the translation report was approved by the Translation Unit of the EORTC Quality of Life Department, the questionnaire was ready for field testing (ie, psychometric evaluation).

During the field testing phase, the questionnaire was completed by 170 patients, aged 39 to 82 years (mean

TABLE 1. Basic demographic characteristics of included subjects ($n = 170$).

Characteristics	No. of patients	%
Sex		
Male	146	85.9
Female	24	14.1
Place of living		
Urban	108	63.5
Rural	62	36.5
Living with		
Alone	22	12.9
Partner	137	80.6
Other	11	6.5
Level of education		
No formal education	9	5.2
Elementary school	43	25.3
Secondary school	71	41.8
High school/faculty	47	27.6
Occupational status		
Employed	28	16.5
Unemployed	21	12.4
Retired	121	71.1
Laryngectomy		
Total	87	51.2
Partial	83	48.8
Stage		
I	39	22.9
II	26	15.3
III	68	40
IV	37	21.8

[SD], 62 years), 7 (8.1) who underwent laryngectomy for laryngeal carcinoma (Table 1). The time between the patients' laryngectomy and this study ranged from 1 month to 18 years (mean, 3.25 ± 3.05 years), with 39 (22.9%) of those included being 5-year survivors. Of all included subjects, 98 (57.6%) also had adjuvant radiation therapy (RT), 8 (4.7%) also had systemic chemotherapy, and 18 (10.5%) also had both.

As a sample of convenience, all subjects were recruited during prescheduled assessments while visiting a doctor for follow-up care. The participation was on a voluntary basis and the only exclusion criterion was inability to read and/or write Serbian. All included patients provided written informed consent. Ten patients (7 men) who were eligible for the study declined to participate. One man reported significant pain and refused to participate, whereas others commented that the questionnaires were long and time-consuming. These nonparticipating patients did not differ significantly from the participants in clinical and demographic variables.

The EORTC QLQ-C30 version 3.0 was also completed.¹² The EORTC QLQ-C30 has the following multi-item function/symptom scales and single items: global health/QOL, physical, role, emotional, cognitive, and social functioning, fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea, and financial difficulties. Scores for multi-item functional or symptom scales and for single items were calculated by linear transformation of raw scores into a 0 to 100 score. Scores of 100 represent the best outcomes on the

TABLE 2. Descriptive statistics for the Quality of Life Questionnaire-Core 43-questions ($n = 170$).

QLQ-H&N43 score	Missing (%)	Minimum	Maximum	Mean	SD	Floor/ceiling (%)
Pain	2	0	83.33	12.90	18.88	47.6/0
Swallowing	2	0	100	11.61	18.98	51.2/0.6
Senses problems	2	0	100	21.73	28.06	52.9/2.4
Speech problems	7	0	100	40.41	29.67	9.4/4.7
Social eating	2	0	100	11.76	18.83	58.2/0.6
Sexuality	6	0	100	30.69	33.13	38.2/7.6
Problems with teeth	4	0	88.89	20.68	23.24	36.5/0
Dry mouth/sticky saliva	3	0	100	28.34	24.92	26.5/2.4
Body image	4	0	100	18.14	24.75	46.5/1.2
Shoulder problems	3	0	100	18.56	25.24	50/2.9
Skin problems	2	0	77.78	9.06	14.87	59.4/0
Anxiety	2	0	100	37.50	28.82	21.8/5.3
Social contact	3	0	100	21.36	32.57	61.8/8.8
Opening mouth	3	0	100	8.38	21.56	82.9/1.8
Coughing	2	0	100	29.37	30.04	40.6/5.9
Lymphedema	3	0	100	14.17	26.99	71.8/4.7
Weight loss	2	0	100	13.10	25.54	72.9/4.1
Problems wounds healing	2	0	100	8.53	21.89	88.8/3.5
Neurological problems	2	0	100	16.67	24.47	61.2/2.4

Abbreviation: QLQ-H&N43, Quality of Life Questionnaire-Core 43-questions.

QLQ-C30 functioning scales and the worst outcomes on the QLQ-C30 symptom scales and the QLQ-H&N43 scales. The Serbian version was provided by the EORTC group.

The study protocol was in keeping with the tenets of the Declaration of Helsinki and its later amendments and was approved by the Ethics Committee of the Faculty of Medicine, University of Belgrade.

Data analysis

The distribution of mean and SD values, percentage of missing values, and floor and ceiling effects were calculated for all QLQ-H&N43 scales. A significant floor and ceiling effect was defined as having >15% of participants scoring at the lowest and highest possible score, respectively.¹³

Internal consistency reliability was assessed for the proposed scales with 3 to 5 items by calculating Cronbach's α , and values of $\alpha \geq 0.7$ were considered acceptable.¹⁴ Additionally, corrected item-total correlations were estimated for each item in a scale. It was assumed that an item should be substantially correlated to the underlying concept measured, indicated by correlations of ≥ 0.4 between the score for each item and its subscale total.¹⁴

In order to determine whether the QLQ-H&N43 scales assess different concepts, correlations among their scores were computed using Spearman's ρ correlation coefficient. A scale was considered redundant if $p > .7$. Furthermore, correlations between the QLQ-C30 and QLQ-H&N43 scales were also estimated (p). It was hypothesized that correlations between the QLQ-C30 functioning scales (ie, global health/QOL, physical, role, emotional, cognitive, and social functioning) and the QLQ-H&N43 should be negative, indicating that the QLQ-C30 functioning domains would be more impaired with more QLQ-H&N43 symptoms present. Finally, to support its

discriminant validity, the QLQ-H&N43 scales should not be highly correlated (ie, $p > .5$) with the QLQ-C30 symptoms scales (ie, fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, diarrhea), thus indicating that the scales indeed measure different concepts.

Construct validity was determined utilizing the known-groups method.^{15,16} The known-groups method compares scores from a scale across groups assumed to differ in the health construct being investigated. In the present study, it was hypothesized that: (1) the subjects with partial laryngectomy would report lower QLQ-H&N43 scores than the subjects with total laryngectomy; (2) the subjects with adjuvant therapy (adjuvant RT, systemic chemotherapy, or both) would report higher QLQ-H&N43 scores than the subjects with laryngectomy only; and (3) 5-year survivors would report lower QLQ-H&N43 scores than patients in active treatment. Differences between the specified groups should be in the proposed fashion, but not necessarily statistically significant. Considering that the QLQ-H&N43 is not a specific questionnaire for patients with laryngectomy and that it is unrealistic to expect that all scores would be in the predicted fashion, the following specific assumptions were made. The subjects with total laryngectomy would have substantially higher scores across the QLQ-H&N swallowing, QLQ-H&N speech problems, QLQ-H&N social eating, QLQ-H&N body image, QLQ-H&N social contact, and QLQ-H&N coughing scales, because these domains could be more substantially affected in total than in partial laryngectomy.¹⁷ Furthermore, it was assumed that subjects with adjuvant therapy would have substantially higher scores across the QLQ-H&N senses problems, QLQ-H&N dry mouth/sticky saliva, QLQ-H&N skin problems, QLQ-H&N lymphedema, and QLQ-H&N weight loss scales, because these domains could be more affected by side effects of RT or systemic chemotherapy. Finally, it was

TABLE 3. Item-scale correlations and internal consistency reliability of the Quality of Life Questionnaire-Core 43-questions multiple-item scales ($n = 170$).

QLQ-H&N43 scale Cronbach's α	H&N item #	Corrected item-total correlation	Cronbach's α if item deleted
Pain 0.83	31 "pain in your mouth"	0.70	0.75
	32 "pain in your jaw"	0.63	0.79
	33 "soreness in your mouth"	0.58	0.81
	34 "pain in your throat"	0.70	0.75
Swallowing 0.86	35 "problems swallowing liquids"	0.79	0.77
	36 "problems swallowing pureed food"	0.78	0.80
	37 "problems swallowing solid food"	0.71	0.81
	38 "choked when swallowing"	0.56	0.87
Speech problems 0.86	47 "problems with hoarseness"	0.35	0.91
	55 "problems talking to other people"	0.73	0.82
	56 "problems talking on the telephone"	0.81	0.79
	57 "problems talking in a noisy environment"	0.81	0.80
	58 "problems speaking clearly"	0.74	0.82
Social eating 0.77	51 "problems eating"	0.41	0.78
	52 "problems eating in front of your family"	0.61	0.69
	53 "problems eating in front of other people"	0.72	0.61
	54 "problems enjoying your meals"	0.55	0.71
Problems with teeth 0.65	39 "problems with your teeth"	0.57	0.39
	40 "problems because of losing some teeth"	0.60	0.33
	73 "problems chewing"	0.25	0.77
Skin problems 0.52	65 "skin problems (eg, itchy, dry)"	0.33	0.60
	66 "rash"	0.43	0.40
	67 "skin changed color"	0.39	0.35
Body image 0.82	48 "problems with your appearance"	0.72	0.68
	49 "felt less physically attractive"	0.66	0.75
	50 "felt dissatisfied with your body"	0.63	0.79

Abbreviation: QLQ-H&N43, Quality of Life Questionnaire-Core 43-questions.

assumed that 5-year survivors would have lower scores across the QLQ-H&N pain, QLQ-H&N speech problems, QLQ-H&N body image, QLQ-H&N anxiety, QLQ-H&N social eating, and QLQ-H&N weight loss scales, considering they are expected to adjust to the condition and better cope with it over time.

The t test was used to assess differences between the groups and Cohen's d effect size was calculated for eventual significant differences. Effect size should be interpreted as: 0.2 small, 0.5 moderate, or 0.8 large.¹⁶

RESULTS

The percentages of missing values per item ranged between 1% and 4%, whereas per scale were between 2% and 7% (Table 2). All scales had floor effects >15%, except for the QLQ-H&N speech problems scale (9.4%). No substantial ceiling effects were found across the scales.

Cronbach's α of above 0.7 was found for 5 of 7 QLQ-H&N43 multi-item scales, namely the QLQ-H&N pain (0.83), QLQ-H&N swallowing (0.86), QLQ-H&N speech problems (0.86), QLQ-H&N social eating (0.77), and QLQ-H&N body image (0.82) scales (Table 3). Within these scales, only item #47 ("problems with hoarseness") had a low corrected item-total correlation (0.35) and removing this item from the QLQ-H&N speech problems scale increased Cronbach's α to 0.91. In the QLQ-H&N problems with teeth scale ($\alpha = 0.65$), item #73 ("problems chewing") had a low corrected item-total correlation

(0.25). In the QLQ-H&N skin problems scale ($\alpha = 0.52$), item #65 ("skin problems [eg, itchy, dry]") and item #67 ("skin changed color"), both had low corrected item-total correlations (0.33 and 0.39, respectively).

All QLQ-H&N43 scale intercorrelations were between 0.02 and 0.67, and thus none of the scales was considered redundant (Table 4). All of the QLQ-C30 functioning scales (ie, global health/QOL, physical, role, emotional, cognitive, and social) were negatively correlated with the QLQ-H&N43 scales, ranging from -0.05 to -0.58 . In terms of the QLQ-C30 symptom scales (ie, the fatigue, nausea and vomiting, pain, dyspnea, insomnia, appetite loss, constipation, and diarrhea), only the QLQ-C30 pain scale was highly correlated with one of the QLQ-H&N43 scales, namely the QLQ-H&N social eating scale ($p = .51$; Table 5).

In terms of the known-group analyses (Table 6), the subjects with total laryngectomy had higher scores than subjects with partial laryngectomy in 16 of 19 QLQ-H&N43 scales. As specifically predicted, the QLQ-H&N speech problems, QLQ-H&N social eating, QLQ-H&N body image, QLQ-H&N social contact, and QLQ-H&N coughing scores were significantly higher in the subjects with total than partial laryngectomy. However, the same prediction for the QLQ-H&N swallowing scale was not observed. The subjects with adjuvant therapy had higher scores than the subjects with laryngectomy only in 17 of 19 QLQ-H&N43 scales, including significantly higher for the QLQ-H&N senses problems, QLQ-H&N dry mouth/sticky saliva, QLQ-H&N skin problems, QLQ-H&N

TABLE 4. Correlations among the Quality of Life Questionnaire-Core 43-questions scales/items.

QLQ-H&N43 score	1	2	3	4	5	6	7	7	9	10	11	12	13	14	15	16	17	18	
1. Pain	-																		
2. Swallowing	0.50																		
3. Senses problems	0.31	0.33																	
4. Speech problems	0.22	0.46	0.50																
5. Social eating	0.42	0.56	0.47	0.63															
6. Sexuality	0.29	0.30	0.44	0.50	0.51														
7. Problems with teeth	0.33	0.27	0.48	0.42	0.47	0.31													
8. Dry mouth/sticky saliva	0.52	0.31	0.22	0.16	0.27	0.27	0.23												
9. Body image	0.28	0.39	0.55	0.63	0.58	0.56	0.45	0.20											
10. Shoulder problems	0.31	0.18	0.17	0.26	0.39	0.28	0.38	0.24	0.28										
11. Skin problems	0.24	0.31	0.22	0.18	0.38	0.21	0.22	0.20	0.27	0.30									
12. Anxiety	0.28	0.30	0.17	0.41	0.40	0.37	0.30	0.12	0.36	0.39	0.31								
13. Social contact	0.24	0.35	0.54	0.67	0.61	0.50	0.45	0.11	0.65	0.29	0.27	0.30							
14. Opening mouth	0.34	0.33	0.23	0.16	0.28	0.23	0.31	0.29	0.24	0.25	0.20	0.18	0.07						
15. Coughing	0.25	0.36	0.25	0.34	0.34	0.34	0.33	0.37	0.32	0.19	0.14	0.12	0.30	0.21					
16. Lymphedema	0.32	0.27	0.23	0.25	0.38	0.19	0.24	0.28	0.17	0.33	0.39	0.18	0.31	0.13	0.25				
17. Weight loss	0.10	0.28	0.20	0.42	0.49	0.30	0.35	0.02	0.40	0.32	0.24	0.26	0.43	0.15	0.22	0.30			
18. Problems wounds healing	0.02	0.20	0.18	0.25	0.28	0.21	0.23	0.03	0.32	0.26	0.33	0.35	0.21	0.29	0.21	0.18	0.33		
19. Neurological problems	0.27	0.22	0.13	0.10	0.17	0.09	0.24	0.20	0.13	0.30	0.34	0.18	0.20	0.05	0.18	0.09	0.10	0.13	

Abbreviation: QLQ-H&N43, Quality of Life Questionnaire-Core 43-questions.

TABLE 5. Correlations among the QLQ-C30 and QLQ-H&N43 scales/items.

QLQ-H&N43 score	QLQ-C30 score														
	Global health/QoL	Physical functioning	Role functioning	Emotional functioning	Cognitive functioning	Social functioning	Fatigue	Nausea & vomiting	Pain	Dyspnea	Insomnia	Appetite loss	Constipation	Diarrhea	Financial difficulties
Pain	-0.40	-0.28	-0.22	-0.32	-0.33	-0.22	0.30	0.37	0.34	0.35	0.20	0.34	0.12	0.17	0.20
Swallowing	-0.38	-0.31	-0.35	-0.36	-0.37	-0.32	0.31	0.29	0.32	0.43	0.22	0.28	0.21	0.24	0.20
Senses problems	-0.24	-0.34	-0.34	-0.30	-0.13	-0.39	0.23	0.22	0.26	0.19	0.13	0.22	0.11	0.07	0.35
Speech problems	-0.31	-0.47	-0.41	-0.53	-0.28	-0.50	0.49	0.25	0.39	0.38	0.24	0.22	0.10	0.24	0.35
Social eating	-0.41	-0.52	-0.44	-0.43	-0.39	-0.40	0.43	0.36	0.51	0.33	0.28	0.37	0.16	0.31	0.37
Sexuality	-0.20	-0.31	-0.19	-0.36	-0.25	-0.37	0.24	0.16	0.41	0.26	0.07	0.28	0.21	0.36	0.33
Problems with teeth	-0.25	-0.37	-0.42	-0.32	-0.25	-0.31	0.30	0.36	0.30	0.21	0.19	0.29	0.08	0.23	0.41
Dry mouth/sticky saliva	-0.43	-0.27	-0.16	-0.30	-0.36	-0.22	0.29	0.15	0.22	0.30	0.20	0.21	0.18	0.12	0.17
Body image	-0.24	-0.49	-0.47	-0.41	-0.21	-0.55	0.39	0.24	0.47	0.36	0.14	0.31	0.27	0.21	0.41
Shoulder problems	-0.36	-0.40	-0.25	-0.32	-0.30	-0.27	0.33	0.32	0.35	0.35	0.23	0.08	0.16	0.26	0.27
Skin problems	-0.26	-0.28	-0.28	-0.16	-0.22	-0.21	0.28	0.29	0.35	0.21	0.23	0.24	0.17	0.14	0.08
Anxiety	-0.37	-0.38	-0.30	-0.58	-0.30	-0.41	0.29	0.31	0.31	0.23	0.18	0.18	0.22	0.25	0.21
Social contact	-0.32	-0.40	-0.49	-0.34	-0.25	-0.45	0.36	0.23	0.31	0.30	0.11	0.25	0.09	0.14	0.31
Opening mouth	-0.20	-0.25	-0.05	-0.26	-0.35	-0.19	0.15	0.14	0.24	0.22	0.09	0.24	0.28	0.18	0.18
Coughing	-0.35	-0.34	-0.23	-0.27	-0.31	-0.23	0.33	0.47	0.26	0.34	0.18	0.27	0.27	0.16	0.19
Lymphedema	-0.30	-0.23	-0.25	-0.20	-0.15	-0.16	0.16	0.19	0.27	0.22	0.07	0.26	0.07	0.03	0.06
Weight loss	-0.26	-0.33	-0.27	-0.27	-0.28	-0.40	0.32	0.08	0.28	0.24	0.08	0.28	0.17	0.18	0.35
Problems wounds healing	-0.23	-0.28	-0.25	-0.32	-0.36	-0.36	0.19	0.33	0.22	0.24	0.06	0.31	0.28	0.18	0.26
Neurological problems	-0.23	-0.27	-0.17	-0.23	-0.27	-0.05	0.26	0.21	0.22	0.21	0.16	0.05	0.14	0.16	0.03

Abbreviations: QLQ-H&N43, Quality of Life Questionnaire-Core 43-questions; QLQ-C30, Quality of Life Questionnaire-Core 30-questions; QoL, quality of life.

TABLE 6. The Quality of Life Questionnaire-Core 43-questions scores according to laryngectomy type, adjuvante therapy, and 5-year survival.

QLQ-H&N43 score	Laryngectomy			Adjuvant therapy			5-y survival		
	Category, no. of patients	Mean (SD)	t-value,* d [†]	Category, no. of patients	Mean (SD)	t-value,* d [†]	Category, no. of patients	Mean (SD)	t-value,* d [†]
Pain	Total, 85 Partial, 83	12.65 (17.89) 13.15 (19.95)	-0.17, -0.03	Yes, 104 No, 64	13.62 (17.45) 11.72 (21.08)	0.63, 0.09	No, 131 Yes, 37	10.88 (16.27) 20.05 (25.11)	-2.65,* 0.49 -0.28, -0.05
Swallowing	Total, 85 Partial, 83	11.18 (19.01) 12.05 (19.05)	-0.29, -0.06	Yes, 104 No, 64	12.50 (19.67) 10.16 (17.84)	0.77, 0.12	Yes, 37 No, 131	11.39 (18.11) 12.39 (22.01)	
Senses problems	Total, 85 Partial, 83	35.69 (29.45) 7.43 (17.51)	7.56,* 1.17	Yes, 104 No, 64	30.93 (29.23) 6.77 (17.99)	5.95,* 0.94	No, 131 Yes, 37	23.16 (29.13) 16.67 (23.57)	1.24, 0.24
Speech problems	Total, 80 Partial, 83	51.17 (31.19) 30.04 (24.09)	4.85,* 0.76	Yes, 101 No, 62	43.89 (31.44) 34.73 (25.78)	1.93, 0.31	No, 127 Yes, 36	44.67 (29.86) 25.37 (23.78)	3.57,* 0.67
Social eating	Total, 85 Partial, 83	16.27 (22.00) 7.13 (13.53)	3.23,* 0.49	Yes, 104 No, 64	15.71 (21.16) 5.34 (11.81)	3.58,* 0.57	No, 131 Yes, 37	12.53 (19.59) 9.01 (15.76)	1.01, 0.19
Sexuality	Total, 82 Partial, 83	39.02 (37.42) 22.36 (25.82)	3.31,* 0.52	Yes, 102 No, 62	33.82 (33.53) 25.54 (32.04)	1.56, 0.25	No, 129 Yes, 35	35.01 (34.22) 14.76 (22.78)	3.30,* 0.63
Problems with teeth	Total, 83 Partial, 83	26.10 (23.82) 15.26 (21.44)	3.08,* 0.48	Yes, 103 No, 63	26.43 (24.45) 11.29 (17.56)	4.28,* 0.68	No, 129 Yes, 37	21.71 (23.68) 17.12 (21.53)	1.06, 0.20
Dry mouth/sticky saliva	Total, 84 Partial, 83	28.77 (25.39) 27.91 (24.57)	0.22, 0.20	Yes, 104 No, 63	31.41 (25.68) 23.28 (22.90)	2.06,* 0.33	No, 130 Yes, 37	29.36 (24.53) 24.77 (26.24)	0.98, 0.18
Body image	Total, 84 Partial, 82	26.85 (28.56) 9.21 (15.88)	4.90,* 0.76	Yes, 103 No, 63	22.44 (27.61) 11.11 (17.16)	2.92,* 0.47	No, 129 Yes, 37	20.33 (25.57) 10.51 (20.10)	2.15,* 0.40
Shoulder problems	Total, 84 Partial, 83	20.44 (28.02) 16.67 (22.08)	0.96, 0.15	Yes, 103 No, 64	20.55 (26.52) 15.36 (22.86)	1.29, 0.21	No, 130 Yes, 37	19.62 (26.04) 14.86 (22.14)	1.01, 0.19
Skin problems	Total, 85 Partial, 83	9.28 (14.32) 8.84 (15.49)	0.19, 0.03	Yes, 104 No, 64	12.07 (17.17) 4.17 (8.04)	3.45,* 0.55	No, 131 Yes, 37	9.58 (15.76) 7.21 (11.17)	0.86, 0.19
Anxiety	Total, 85 Partial, 83	37.25 (30.39) 37.75 (27.31)	-0.11, -0.02	Yes, 104 No, 64	36.06 (29.84) 39.84 (27.15)	-0.82, -0.13	No, 131 Yes, 37	40.20 (28.18) 27.93 (29.41)	2.31,* 0.43
Social contact	Total, 84 Partial, 83	34.13 (36.94) 8.43 (20.72)	5.53,* 0.86	Yes, 103 No, 64	28.48 (36.57) 9.90 (20.29)	3.72,* 0.59	No, 131 Yes, 36	24.17 (33.60) 11.11 (26.42)	2.16,* 0.41
Opening mouth	Total, 84 Partial, 83	9.52 (24.03) 7.23 (18.80)	0.68, 0.11	Yes, 104 No, 63	8.33 (21.15) 8.47 (22.37)	-0.03, -0.09	No, 130 Yes, 37	9.23 (23.11) 5.41 (14.72)	0.95, 0.20
Coughing	Total, 85 Partial, 83	34.51 (31.89) 24.10 (27.20)	2.27,* 0.35	Yes, 104 No, 64	31.73 (30.59) 25.52 (28.93)	1.30, 0.10	No, 131 Yes, 37	30.53 (30.39) 25.23 (28.76)	0.95, 0.18
Lymphedema	Total, 84 Partial, 83	16.67 (29.94) 11.65 (23.53)	1.20, 0.19	Yes, 103 No, 64	20.06 (31.08) 4.69 (14.38)	3.71,* 0.59	No, 130 Yes, 37	15.38 (28.50) 9.91 (20.58)	1.09, 0.22
Weight loss	Total, 85 Partial, 83	18.82 (29.74) 7.23 (18.80)	3.01,* 0.46	Yes, 104 No, 64	17.31 (28.62) 6.25 (17.69)	2.77,* 0.44	No, 131 Yes, 37	15.27 (26.87) 5.41 (18.44)	2.09,* 0.38
Problems wounds healing	Total, 85 Partial, 83	10.59 (25.82) 6.43 (16.83)	1.23, 0.19	Yes, 104 No, 64	10.90 (26.04) 4.69 (11.67)	1.79, 0.31	No, 131 Yes, 37	10.69 (24.20) 0.90 (5.48)	2.44,* 0.45
Neurological problems	Total, 85 Partial, 83	17.25 (24.99) 16.06 (24.06)	0.31, 0.05	Yes, 104 No, 64	17.95 (24.98) 14.58 (23.66)	0.86, 0.14	No, 131 Yes, 37	16.79 (24.93) 16.22 (23.06)	0.13, 0.02

* *p* < .05.
† Cohen's *d* effect size.

lymphedema, and QLQ-H&N weight loss scales. Additionally, the QLQ-H&N social eating, QLQ-H&N problems with teeth, QLQ-H&N body image, and QLQ-H&N social contact scores were also statistically higher in the subjects with adjuvant therapy than subjects who received no adjuvant treatment (Table 6). Finally, the 5-year survivors had lower scores in 17 of 19 QLQ-H&N43 scales. The QLQ-H&N speech problems, QLQ-H&N body image, QLQ-H&N anxiety, and QLQ-H&N weight loss scores were significantly lower among the subjects on treatment than 5-year survivors, as predicted. The QLQ-H&N social eating score was also lower, but not statistically significant. Contrary to our hypothesis, the QLQ-H&N pain score was significantly higher among 5-year survivors than the subjects on treatment. Two more scores also significantly differed between the groups; the QLQ-H&N sexuality and QLQ-H&N social contact (Table 6).

DISCUSSION

As a whole, the amount of missing data for the items in the module was small, but it was relatively high (up to 7%) for 2 scales, namely the QLQ-H&N speech problems and QLQ-H&N sexuality. This finding supports the notion that the items in its assumed scales have good feasibility and acceptability in this particular group. Additionally, negligible ceiling effects were found across all scales. However, relevant floor effects (above 15%) were detected for all except for the QLQ-H&N speech problems scale, indicating that the percentage of individuals with the best results, or indeed the absence of symptoms, was high. Although this may be an indication of low discriminative power of the module for detecting favorable outcomes, it is also possible that our subjects minimized their symptoms or were symptom-free at the time of assessment, considering that almost 23% of included subjects in the study were 5-year survivors. Having only a small number of items per scale (1–4) may have also contributed to the high floor effect. Similarly, previous studies with the QLQ-H&N35 also reported relevant floor effects.^{18,19}

The internal consistency of the Serbian version was found to be adequate for a majority of the scales. The 2 scales with low internal consistency were the QLQ-H&N problems with teeth and the QLQ-H&N skin problems, both of which consisted of 3 items. These findings disagree with the original study that showed adequate internal consistency for all proposed scales (Cronbach's α ranged between 0.78 and 0.86 in that study).⁹ Considering item-level statistics, 4 items were problematic: item #47 ("problems with hoarseness") in the QLQ-H&N speech problems scale, item #73 ("problems chewing") in the QLQ-H&N problems with teeth, and item #65 ("skin problems [eg, itchy, dry]") and item #67 ("skin changed color") in the QLQ-H&N skin problems. These items showed low correlations with their respective scale scores, which might indicate that they measure different concepts from other items in their scale. On the contrary, none of the items was problematic in the original study.⁹ Considering that this study included only laryngectomized patients, future research by the EORTC group should focus on evaluating item-scale correlations across groups

of patients with different carcinomas in subsequent revisions of the QLQ-H&N43.

Additional data showed that all QLQ-H&N43 scale intercorrelations were lower than 0.7, consistent with our hypothesis that all scales measure different constructs and that none of the scales are redundant. Moreover, all QLQ-H&N43 scales correlated adequately with all QLQ-C30 functioning scales, suggesting that the QLQ-C30 functioning domains would be more impaired if QLQ-H&N43 symptoms are present. The small to moderate correlations found between the QLQ-C30 and QLQ-H&N43 scales suggest that the 2 measures captured different QOL domains, providing support for their discriminant validity. One exception to this was the moderately high correlation between the QLQ-H&N social eating scale and the QLQ-C30 pain scale. Upon inspecting the items in these scales, the correlation does not seem to be due to any overlap or similarities in item content, and thus the correlation possibly indicated how the presence of pain may be associated with more eating difficulties. Overall, the above findings suggest that the entire module could be used with the core measure for QOL assessments in patients with laryngectomy, which is in line with the EORTC guidelines.¹²

With regard to the measures construct validity, this study showed that specific QLQ-H&N43 scales scores were higher for particular groups; namely, those who underwent total laryngectomy, those who received adjuvant therapy, and those who are still in active treatment. Specifically, we anticipated significant group differences for 17 comparisons, and 13 of these were supported. The results of the known-group comparisons showed that 3 scales, namely QLQ-H&N43 body image, QLQ-H&N43 social contact, and QLQ-H&N43 weight loss, demonstrated significant group differences across all 3 group comparisons. Another 10 scales demonstrated significant differences in the predicted fashion for at least 1 of the 3 group comparisons. In addition, it was also observed that patients with total laryngectomy and those who had adjuvant therapy had significantly higher scores in the QLQ-H&N problems with teeth scale than patients with partial laryngectomy or those who did not have adjuvant therapy. However, the QLQ-H&N43 pain and QLQ-H&N43 swallowing did not demonstrate any of the predicted group differences. According to our results, the QLQ-H&N swallowing scale might not be able to differentiate swallowing difficulties in patients with total versus partial laryngectomy, although it is expected that patients with total laryngectomy would have more swallowing difficulties. Another explanation would be that the patients in our study did not differ in terms of their swallowing problems; hence the QLQ-H&N43 could not find any differences. The same applies for the QLQ-H&N pain scale that might not differentiate different levels of pain in the mouth, jaw, and throat in patients with total and patients with partial laryngectomy. Nevertheless, the QLQ-H&N pain scale scores were significantly higher in 5-year survivors than among patients on treatment, which is also unexpected and might indicate that this scale is not valid for patients with laryngectomy. Another explanation could be a phenomenon that has been called the "satisfaction paradox" – patients with partial laryngectomy may tend to expect better outcomes than patient who underwent

total laryngectomy, resulting in overestimation of current problems.²⁰ Considering these findings, construct validity for the QLQ-H&N43 should be further tested, such as including other groups of patients known to differ in health status or by the means of assessing convergent and discriminant validity.

Some limitations of this study need to be acknowledged. First, we have included only subjects with laryngeal carcinoma who underwent laryngectomy with/without adjuvant therapy. As such, patients with other head and neck cancers and those who underwent other treatments need to be included in order to provide greater variability and validity in the module scores. For example, the QLQ-H&N skin scale might be the most relevant to those patients who are receiving targeted therapy. As the patients in this study had not received any biological treatment, psychometric properties of this scale might be relatively poor. Second, the convergent validity of the newly established scores should be evaluated with well-established measures, such as measures of anxiety levels or body image. Third, test-retest reliability could not be assessed because of budget restrictions of this study. If the module is to be used for follow-up evaluations, its responsiveness and sensitivity to change should be evaluated.

In conclusion, preliminary evidence from patients with laryngectomy suggests that the psychometric properties of the QLQ-H&N43 are acceptable. The module is feasible and easily scored and none of the scales are considered redundant. Five of 7 multiple-item scales in the module have appropriate internal consistency. Preliminary evidence of construct validity was suggested for 14 of 19 scales with significant group differences. Given the validation of a questionnaire is an iterative process, the QLQ-H&N43 will be further field tested in other cancer groups with larger and culturally diverse populations. This study can inform the upcoming EORTC phase IV study, which will test this questionnaire on a larger scale.

REFERENCES

- Acquadro C, Berzon R, Dubois D, et al. Incorporating the patient's perspective into drug development and communication: an ad hoc task force report on the Patient-Reported Outcomes (PRO) Harmonization Group meeting at the Food and Drug Administration, February 16, 2001. *Value Health* 2003;6:522–531.
- Muzzatti B, Annunziata MA. Assessing quality of life in long-term cancer survivors: a review of available tools. *Support Care Cancer* 2013;21:3143–3152.
- Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365–376.
- Velikova G, Coens C, Efficace F, Greimel E, Grønvdol M, Johnson C. Health-related quality of life in EORTC clinical trials – 30 years of progress from methodological developments to making a real impact on oncology practice. *Eur J Cancer Suppl* 2012;10:141–149.
- Bjordal K, Ahlner-Elmqvist M, Tolleson E, et al. Development of a European Organization for Research and Treatment of Cancer (EORTC) questionnaire module to be used in quality of life assessments in head and neck cancer patients. *EORTC Quality of Life Study Group. Acta Oncol* 1994;33:879–885.
- Bjordal K, Hammerlid E, Ahlner-Elmqvist M, et al. Quality of life in head and neck cancer patients: validation of the European Organization for Research and Treatment of Cancer Quality of Life Questionnaire–H&N35. *J Clin Oncol* 1999;17:1008–1019.
- Bjordal K, de Graeff A, Fayers PM, et al. A 12 country field study of the EORTC QLQ-C30 (version 3.0) and the head and neck cancer specific module (EORTC QLQ-H&N35) in head and neck cancer patients. *EORTC Quality of Life Group. Eur J Cancer* 2000;36:1796–1807.
- Singer S, Arraras J, Chie WC, et al. Performance of the EORTC questionnaire for the assessment of quality of life in head and neck cancer patients EORTC QLQ-H&N35: a methodological review. *Qual Life Res* 2013;22:1927–1941.
- Singer S, Araújo C, Arraras JI, et al. Measuring quality of life in patients with head and neck cancer: update of the EORTC QLQ-H&N module, phase III. *Head Neck* 2014. [Epub ahead of print]
- Singer S, Arraras JI, Baumann I, et al. Quality of life in patients with head and neck cancer receiving targeted or multimodal therapy—update of the EORTC QLQ-H&N35, phase I. *Head Neck* 2013;35:1331–1338.
- Koller M, Aaronson NK, Blazeby JM, et al. Translation procedures for standardised quality of life questionnaires: the European Organisation for Research and Treatment of Cancer (EORTC) approach. *Eur J Cancer* 2007;43:1810–1820.
- Fayers PM, Aaronson NK, Bjordal K, et al. The EORTC QLQ-C30 scoring manual. 3rd edition. Brussels, Belgium: European Organisation for Research and Treatment of Cancer; 2001.
- Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol* 2007;60:34–42.
- Nunnally JC, Bernstein IR. *Psychometric theory*. New York, NY: McGraw-Hill; 1994.
- Fayers PM, Machin D. *Quality of life: the assessment, analysis and interpretation of patient-reported outcomes*. 2nd edition. Chichester, NY: John Wiley & Sons; 2007.
- Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.
- Braz DS, Ribas MM, Deditis RA, Nishimoto IN, Barros AP. Quality of life and depression in patients undergoing total and partial laryngectomy. *Clinics (Sao Paulo)* 2005;60:135–142.
- Bower WF, Vlantis AC, Chung TM, Cheung SK, Bjordal K, van Hasselt CA. Quality of life in head and neck cancer patients after surgical resection: translation into Cantonese and validation of the EORTC QLQ-H&N35. *Acta Otolaryngol* 2009;129:779–785.
- Carrillo JF, Ortiz-Toledo MÁ, Salido-Noriega Z, Romero-Ventura NB, Ochoa-Carrillo FJ, Oñate-Ocaña LF. Validation of the Mexican Spanish version of the EORTC QLQ-H&N35 instrument to measure health-related quality of life in patients with head and neck cancers. *Ann Surg Oncol* 2013;20:1417–1426.
- Bindewald J, Herrmann E, Dietz A, et al. Quality of life and voice intelligibility in laryngeal cancer patients—relevance of the “satisfaction paradox” [in German]. *Laryngorhinootologie* 2007;86:426–430.