

SYSTEMATIC REVIEW

Open Access



Systematic review and standardized comparison of health-related quality-of-life instruments available for patients with oral cavity cancer and oropharyngeal cancer

Naira Figueiredo Deana^{1,2,3}, Carolina Ulloa⁴, Francisco Ceballos^{5,6,7}, Nilton Alves^{6,8,9}, Rodrigo Mariño^{1,10,11}, Wilfredo Alejandro González-Arriagada¹² and Carlos Zaror^{1,2*}

Abstract

Background To identify HRQoL instruments available for oral cavity cancer (OCC) or oropharyngeal cancer (OC) patients and to summarize the evidence on conceptual model measurement, psychometric and administration properties.

Materials and methods We searched MEDLINE, CENTRAL, EMBASE, and LILACS until April 2023. We included studies evaluating the concept model measurement, psychometric and administration properties of instruments measuring HRQoL in OCC or OC patients. Two researchers assessed each instrument independently, using the Evaluating Measures of Patient-Reported Outcomes (EMPRO) tool. We calculated the overall and seven attribute-specific scores (range 0–100) using the conceptual and measurement model, reliability, validity, responsiveness, interpretability, burden, and alternative forms.

Results We included seven instruments across 42 articles. The overall EMPRO score of these instruments varied between 84.1 and 29.3, with five questionnaires exceeding the threshold score of 50.0. The domains of conceptual and measurement model and interpretability performed best, while responsiveness exhibited the lowest performance.

Conclusions The findings support the use of EORTC QLQ-H&N43, EORTC QLQ-C30, and EORTC QLQ-OH15 as suitable HRQoL instruments for both cross-sectional or longitudinal studies to measure physical, symptomatic, psychological, or social impacts in patients with OCC or OC. FACT-H&N and UW-QoL were deemed appropriate for evaluating HRQoL in cross-sectional studies in these patients.

Keywords EMPRO tool, Mouth neoplasms, Head and Neck cancer, Health-related quality of life, Oropharyngeal neoplasms, Psychometric properties

*Correspondence:

Carlos Zaror
carlos.zaror@ufrontera.cl

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

Introduction

Head and neck cancer (H&NC) ranks as the seventh most common type of cancer worldwide. It comprises laryngeal, pharyngeal (including oropharyngeal), and oral cavity cancers [1]. The oral cavity and oropharynx are responsible for 2–4% of all human cancers and 3.7% of cancer-related deaths [2]. Oral cavity cancer (OCC) affects the oral cavity mucosa, including the tongue, palate, buccal mucosa, floor of the mouth, gingiva, and buccal vestibule [3]. The oropharyngeal cancer (OC) comprises the palatine tonsils, base of the tongue, soft palate, tonsillar pillars and uvula [4]. Generally affecting men over 45 years old, it is related to risk factors such as tobacco, alcohol, betel chewing, diet disturbances, and human papillomavirus (HPV) [5]. An increase in cases among young patients and women over 60 years old has been reported in recent decades [6].

Most cases of OCC and OC are diagnosed at advanced stages, which significantly affects treatment options and prognosis [7]. As a result, the survival rate following a diagnosis of these cancers is generally low, with a five-year survival rate of about 50% [8]. However, in high-income countries, this survival rate can improve to around 70% [9].

Treatment for OCC and OC is highly individualized, typically involving a combination of surgery, radiotherapy, chemotherapy, and, increasingly, targeted and immunotherapies. Surgery is the mainstay for OCC, while OC are often managed with radiotherapy or chemoradiation, especially in advanced cases. The choice of treatment depends on tumour stage, location, HPV status, and patient factors [10–12].

Chemotherapy and radiotherapy may produce collateral effects, such as oral candidiasis, xerostomia, oral mucositis, dysgeusia, hypofunction of the salivary glands, and osteonecrosis of the mandible [13]. Otherwise, in cases where surgical treatment is necessary, it may cause aesthetic and functional problems in these patients [14, 15], thereby impacting their quality of life [16]. Health-related quality of life (HRQoL) questionnaires [17] assess how people perceive their health related to cancer and its treatment. This includes non-medical aspects such as physical, functional, emotional and mental welfare and non-health-related factors like work, family, friends and others [18]. Evidence shows that patients with OCC and OC experience a lower HRQoL compared with healthy individuals, largely due to the impact of oncotherapy. Additionally, HRQoL further declines when various treatment modalities are combined [19].

Cancer treatment sequelae differ with the location of the tumour and the type of therapy [20–22]; it is, therefore, essential to use an appropriate instrument for every kind of tumour. The EORTC QLQ-H&N35, UW-QOL, and FACT-HN are among the most commonly

utilized instruments for assessing HRQoL in OCC and OC patients [23]; however, they are not specific for the assessment of OCC and OC since they were developed for other types of cancers. Using instruments in populations that were not created or validated without knowing their performance could overestimate or underestimate the actual impact on HRQoL in these patients.

Accordingly, the research question to be answered is as follows: To what extent are the metrics of each instrument robust and suitable for assessing oral health-related quality of life in OCC and OC patients? Guided by this question, our study aimed to identify and summarize the HRQoL questionnaires available for OCC and OC patients, focusing on their conceptual frameworks, measurement models, psychometric properties, interpretability, and administration features. By comparing and evaluating the strengths and weaknesses of each instrument, we sought to provide evidence that supports the selection of the most appropriate assessment tool for both clinical practice and research applications.

Methods

Protocol

We reported this systematic review using the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines [24, 25] (Online Resource 1) and registered it in PROSPERO (CRD42020220089).

Eligibility criteria

We included qualitative, observational, and experimental research in English, Spanish or Portuguese that provides details on the conceptual and measurement model, psychometric properties (i.e. reliability, validity and responsiveness), interpretability, and administration (i.e. administration burden and alternative modes of administration) of HRQoL instruments in OCC or OC patients (after diagnosis, radiotherapy, chemotherapy, surgical or combined treatment). We also included studies reporting the development of instruments that had initially not been identified in the search, regardless of the type of cancer.

We excluded reports evaluating patient-reported outcomes (PRO) other than the HRQoL, studies reporting data for different types of cancer without specifying results for OCC and OC, and studies analysing other diseases.

Information sources and search

We systematically searched until April 2023 in MEDLINE, CENTRAL, EMBASE, and LILACS. In addition, we examined the reference list of included articles and online databases of PRO instruments: BiblioPRO (www.bibliopro.org) and PROQOLID (<https://www.qolid.org/>). The search strategy used in PubMed

was: (((((((((((questionnaire*) OR instrument*) OR scale*) OR index*) OR batter*) OR inventor*) OR measur*) OR rating*) OR item*) OR Psychometrics) OR "Psychometrics"[Mesh]) OR psychometric*) OR feasibility OR reliability OR validation OR validity OR develop* OR translat* OR Chronbach* AND (((((((("Cancer needs questionnaire") OR ("CNQ-SF-hn") OR ("EORTC QLQ-C30") OR ("EORTC QLQ-H&N43") OR ("EORTC QLQ-OH15") OR ("University of Washington quality of life questionnaire") OR ("UW-QOL") OR ("Functional Assessment of Cancer Therapy") OR ("FACT-H&N") OR ("QOL-OC") OR ("Oral Cancer Quality-of life Questionnaire") AND (((("oral cancer") OR ("oropharyngeal cancer")) OR ("mouth neoplasms"[MeSH Terms])) OR ("Oropharyngeal Neoplasms"[MeSH Terms]))).

Study selection

Two reviewers (C.U, F.C.) selected for eligibility independently titles, abstracts and full-text. A third reviewer (N.A. or C.Z.) resolved discrepancies.

Data collection process

Pairs of trained reviewers with experience in measuring PRO (C.U.-N.F.D., C.Z.-N.F.D.) evaluated each instrument. Reviewers resolved discrepancies through consensus.

Evaluating measures of PRO

The Evaluating Measures of Patient-Reported Outcomes (EMPRO) tool [26] comprises eight attributes: 1. Conceptual and measurement model (7 items); 2. Reliability (8 items); 3. Validity (6 items); 4. Responsiveness (3 items); 5. Interpretability (3 items); 6. Burden (7 items); 7. Alternative modes of administration (2 items); 8. Cross-cultural and linguistic adaptation (3 items). We did not consider the last attribute since it was outside the scope of our study.

Each item is rated on a four-point Likert scale, from 4 (strongly agree) to 1 (strongly disagree), together with a “no information” option (score of 1). In addition, five items allow a rate as “not applicable.” The scores of the first five attributes are used to construct the overall score [26].

Data synthesis

We calculated the overall and attribute-specific scores for each HRQoL instrument when at least 50% of the items were rated. We linearly transformed the mean responses to a range from 0 to 100 (worst to best score). We calculated two sub-scores for the “Reliability” (i.e. internal consistency and reproducibility) and “Burden (i.e. respondent” and administrative) attributes. We chose the higher sub-score for reliability since both elements represent different approaches to examining the same

attribute. For burden, we calculated the mean as the final score since both components evaluate different traits of the same attribute.

We computed the overall score based on the rating of the first five attributes and when a score was registered for at least three of these attributes.

We considered EMPRO scores acceptable if they achieved a minimum of 50 points [26–29]. We included a recommendation scale for the instruments following the criteria of Maratia et al. [30]: Strongly recommended (overall EMPRO score higher than 50.00 in all the attributes), Recommended with provisos or alterations (overall EMPRO score higher than 50.00), and Not recommended (overall EMPRO score lower than 50.00).

Results

After identifying 8,211 references, excluding duplicates, and reviewing titles and abstracts, we analysed 55 full texts. As a result, we eliminated 39 articles: 6 did not evaluate psychometric properties, 15 did not evaluate HRQoL, and 18 did not report data for a population with oral or oropharyngeal cancer. We identified 26 additional articles by manual search in online PRO databases. Finally, we included 42 articles in full-text evaluating seven instruments (Fig. 1). The number of articles for each instrument ranged from 1 to 15, and some reports included data for multiple instruments. Online Resource 2 presents the characteristics of the included articles.

Characteristics of instruments

Table 1 presents the seven instruments applicable to patients with OCC and OC. The instruments were published between 1988 and 2022. Two instruments were developed in English, two in Chinese, and three instruments of the European Organisation for Research and Treatment of Cancer (EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Questionnaire; EORTC QLQ-H&N43: European Organisation for Research and Treatment of Cancer head and neck cancer-specific Quality of Life questionnaire; EORTC QLQ-OH15: European Organisation of Research and Treatment of Cancer, Oral Health Module) were developed in various languages (Dutch, English, French, German, Greek, Hebrew, Italian, Norwegian, Polish and Swedish). Only two instruments, CNQ-SF-hn (Cancer Needs Questionnaire, short form, head and neck cancer-specific) and the QOL-OC (Oral Cancer Quality of Life questionnaire), were not adapted to other languages. The UW-QOL (University of Washington Quality of Life questionnaire) and FACT-H&N (Functional Assessment of Cancer Therapy scale) were developed in English and adapted to other languages. All instruments were developed for self-administration, except EORTC QLQ-OH15, designed for administration

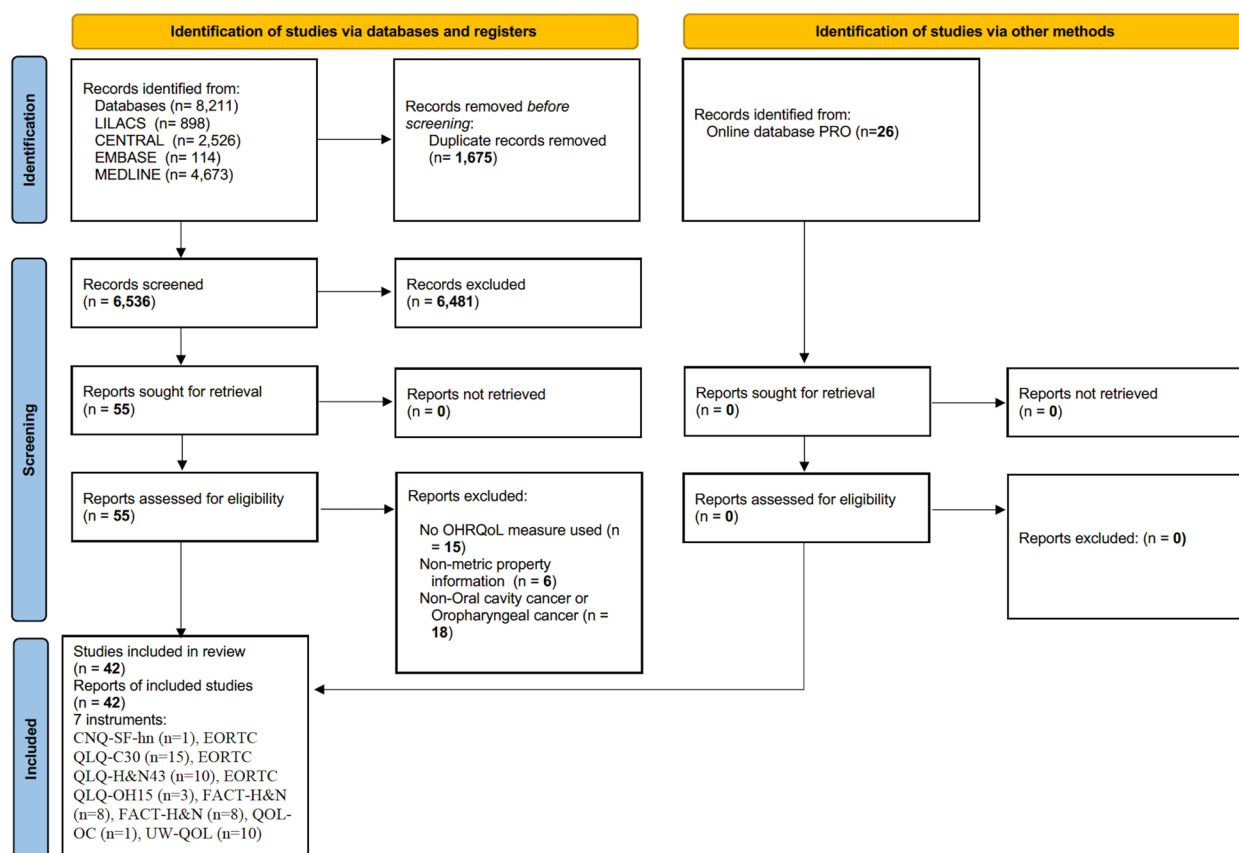


Fig. 1 Flow-chart of the studies and reports included. CNQ-SF-hn: Cancer needs questionnaire, short form, head and neck cancer-specific; EORTC QLQ-C30: European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire; EORTC QLQ-H&N43: European Organization for Research and Treatment of Cancer head and neck cancer specific quality of life questionnaire; EORTC QLQ-OH15: European organization of Research and Treatment of Cancer, Oral Health Module; FACT-H&N (4.0): Functional assessment of cancer therapy scale; QOL-OC: Oral cancer quality-of-life questionnaire; UW-QOL: University of Washington quality of life questionnaire

by interviews. CNQ-SF-hn and QOL-OC are specific for patients with oral cavity cancer; EORTC QLQ- H&N43, FACT-H&N and UW-QOL are specific for H&NC patients; EORTC QLQ-OH15 is a specific instrument for evaluating oral health-related quality of life (OHRQoL) in cancer patients; and EORTC QLQ-C30 is a specific instrument for HRQoL and the perceived health of patients with Neoplasms. EORTC QLQ- H&N43 and EORTC QLQ-OH15 are modules of EORTC which should be applied jointly with EORTC QLQ-C30, a generic instrument designed for any kind of cancer.

Results of the EMPRO ratings

Figure 2 summarizes the overall and attribute-specific scores of the included instruments.

The conceptual and measurement model performed excellently, with all seven instruments scoring above 50.0. The thresholds for this attribute ranged from 59.5 to 97.6, and 3 out of 7 instruments scored higher than 90 (EORTC QLQ-C30, EORTC QLQ-H&N43, EORTC QLQ-OH15).

The Reliability thresholds varied between 25.0 and 62.5 points. EORTC QLQ-H&N43 and the EORTC QLQ-OH15 received the highest scores, followed by EORTC QLQ-C30. The other instruments evaluated scored below 50 for this domain.

Validity was the third-best performing attribute, with 4 out of 7 instruments achieving a score above 50.0. The thresholds ranged from 40.0 for CnQ-SF-hn and QoL-OC and 90.0 for EORTC QLQ-H&N43.

Interpretability achieved the second-best performance. Five instruments received scores higher than 50, with the EORTC QLQ-C30 achieving the highest score (83.3) and the CnQ-SF-hn and the QoL-OC obtaining the lowest (22.2).

Only five instruments reported sufficient data to evaluate Responsiveness; the best score was found for EORTC H&N43 with 94.5, while the lowest was UW-QoL with 27.8.

Regarding ease of use of the instruments, EORTC QLQ-H&N43 and EORTC QLQ-OH15 obtained the highest scores for respondent Burden (77.8 and 72.2, respectively) because they described the skills and time

Table 1 Summarized characteristics of instruments designed or validated for older adults, in alphabetical order

Instrument	Country of development	Purpose of development (population)	Administration mode	Dimensions (no. of items)	Response options	Score (range)	Evaluation	Original and adopted languages	Num-ber of studies evaluated
CNQ-SF-hn (Chinese version)	Taiwan	Head and neck cancer (Specific)	Self-administered	Psychological needs (11), health information needs (7), physical and daily living needs (6), patient care and support needs (6), interpersonal/communication needs (2), head and neck disease needs (4)	36 items, 5-point Likert scale	Score from 0 to 100 (total scale and each subscale)	Higher scores indicate greater need for attention in each domain	Chinese	1
EORTC QLQ-C30	Western Europe, North America, Australia, Japan	HRQoL and perceived health for Neoplasms (Specific)	Self-administration	Physical, role, emotional, cognitive, and social, three symptom scales: fatigue, pain and nausea/vomiting and global health status scale; six single items to assess symptoms: dyspnea, appetite loss, sleep disturbance, constipation, diarrhoea, and financial impact (30)	30 items and 5 functioning subscales	Score from 0 to 100 (all scales and single items)	A high score for a functional scale represents a high/healthy functional level. A high score for the overall state of health/QoL represents a high QoL. A high score for the symptoms/item scale represents a high level of symptomatology/problems	English Azeri Canadian French Dutch (NL) French German Greek Italian Japanese Norwegian Portuguese Swedish	15
EORTC QLQ-H&N43	Belgium, Denmark, Germany, Greece, Italy, Israel, Netherlands, Norway, Serbia, Spain, Sweden, Switzerland, Taiwan, United Kingdom, United States	Head and neck cancer (Specific)	Self-administration	(43)	43 items and symptom subscale	Score from 0 to 100 (all scales and single items)	A high score for a functional scale represents a high/healthy functional level. A high score for the overall state of health/QoL represents a high QoL. A high score for the symptoms/item scale represents a high level of symptomatology/problems	English Azeri Dutch (NL) French German Greek Norwegian Portuguese Swedish	10

Table 1 (continued)

Instrument	Country of development	Purpose of development (population)	Administration mode	Dimensions (no. of items)	Response options	Score (range)	Evaluation	Original and adopted languages	Num-ber of studies evaluated
EORTC QLQ-OH15	10 countries: France, Ger-many, Greece, Israel, Nether-lands, Norway, Poland, Sweden, UK	Oral health and QoL in cancer patients (Specific)	Interview administered	OH-QoL scale (8), 3 single items (sticky saliva/mouth soreness/ sensitivity to food/drink), 2 two-item contingency scales regard-ing use (yes/no) and problems with dentures and reception of (yes/no) and satisfaction with information	4-point Likert scale	Score from 0 to 100 (all scales and single items)	A high score for a functional scale represents a high/healthy functional level. A high score for the overall state of health/QoL represents a high QoL. A high score for the symptoms/item scale represents a high level of symptomatology/problems	Dutch English French German Greek Hebrew Norwegian Polish Sinhalese Swedish	3
FACT-H&N	USA	Head and neck cancer (Specific)	Self-administered	Physical well-being, emotional well-being, functional well-being, head & neck (39)	5-point Likert scale	(0–148)	Higher scores indicate better QoL	English Arabic Malay Urdu (Pakistani)	8

Table 1 (continued)

Instrument	Country of development	Purpose of development (population)	Administration mode	Dimensions (no. of items)	Response options	Score (range)	Evaluation	Original and adopted languages	Num-ber of studies evaluated
QOL-OC	China	Patients with OCC (Specific)	Self-administered	Pain & discomfort, eating, saliva, social contact, eating (29)	4-point Likert Scale (26 items); Yes/No, 1 point (2 items); and open-ended question (non-numerical score)	Score interval (min–max): 28–108	Higher scores indicate worse QoL	Chinese	1
UW-QOL	USA	Head and neck cancer (Specific)	Self-administered	12 items, 12 domains: Pain, appearance, activity, recreation, swallowing, chewing, speech, taste, saliva, mood, anxiety, intimacy and fear of cancer recurrence, and three global quality of life questions. Contains three global questions: one on a 5-point Likert scale which asks about HRQoL in comparison with the month prior to the cancer, and the other two on a 6-point Likert scale, one asking about health and the other about quality of life in general during the "last 7 days"	5-point Likert scale	Scores of 0–100 for each item. Composite QoL score = 100 (max)	Higher scores indicate better QoL	English Indian Portuguese Chinese (Mandarin)	10

OCC Oral cavity cancer, *CNQ-SF-In* Cancer needs questionnaire, short form, head and neck cancer-specific, *EORTC QLQ-C30* European Organization for Research and Treatment of Cancer Quality of Life Core Questionnaire, *EORTC QLQ-H&M43* European Organization for Research and Treatment of Cancer head and neck cancer specific quality of life questionnaire, *EORTC QLQ-OH15* European organization of Research and Treatment of Cancer, Oral Health Module, *FACT-H&N (4.0)* Functional assessment of cancer therapy scale, *QOL-OC* Oral cancer quality-of-life questionnaire, *UW-QOL* University of Washington quality of life questionnaire, *OC* Oral cancer, *HRQoL* Health-related quality of life, *QoL* Quality of life

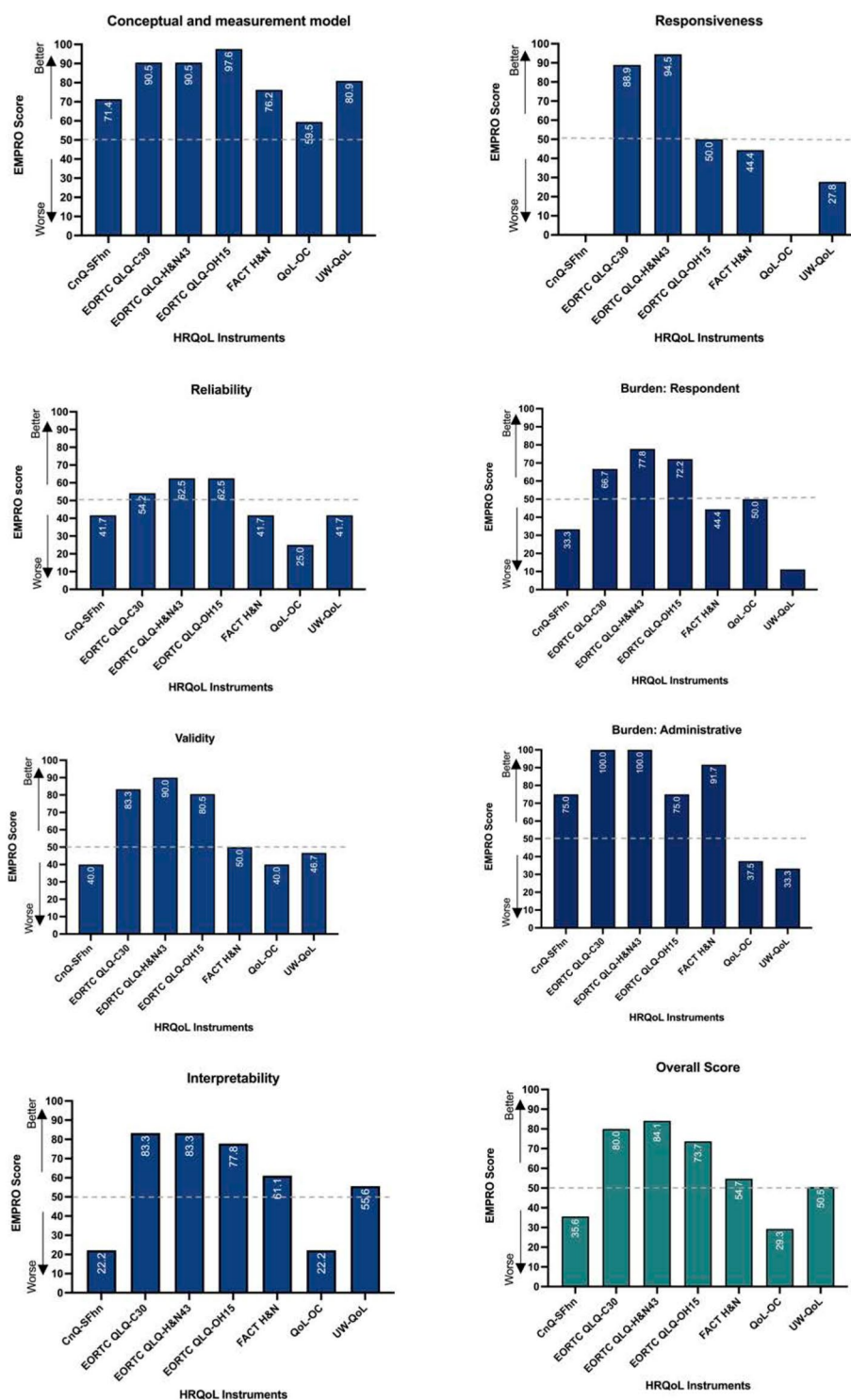


Fig. 2 Overall EMPRO ranking and attribute-specific scores of instruments designed for Oral cavity and Oropharyngeal cancer. The grey line on 50 (half of the theoretical maximum of 100 points) represents the reasonably acceptable cut-off defined for EMPRO scores. CNQ-SF-hn: Cancer needs questionnaire, short form, head and neck cancer-specific; EORTC QLQ-C30: European Organisation for Research and Treatment of Cancer Quality of Life Core Questionnaire; EORTC QLQ-H&N43: European Organisation for Research and Treatment of Cancer head and neck cancer-specific quality of life questionnaire; EORTC QLQ-OH15: European organisation of Research and Treatment of Cancer, Oral Health Module; FACT-H&N (4.0): Functional assessment of cancer therapy scale

required to finish the instrument, its acceptance, and the situations in which it is not suitable for the interviewed. EORTC QLQ-H&N43 and EORTC QLQ-C30 achieved a perfect score (100.0) for questionnaire administration and scoring. This high score reflects the clear description of the resources required for administration, well-defined score calculation methods, and acceptable associated burdens. (Fig. 2).

EORTC QLQ-OH15 and EORTC-H&N43 presented alternative forms of administration. EORTC QLQ-OH15 was developed for interview application but was self-administered in some of the studies analysed. EORTC-H&N43 is the updated version of EORTC-H&N35, and many of the studies included used the older version (Fig. 2).

The instrument with the best EMPRO overall score was EORTC QLQ-H&N43 with 84.1, followed by EORTC QLQ-C30 with 80.0 and EORTC QLQ-OH15 with 73.7 and classified as strongly recommended. FACT H&N and UW-QoL obtained an overall score over 50, meaning their use is recommended with alterations. Finally, CNQ-SF-hn and QoL-OC received an overall score below 50.0; therefore, their use is not recommended. (Fig. 2).

Table 2 shows detailed EMPRO results for any specific criterion and attribute.

Discussion

HRQoL is a predictor of overall survival in patients with OCC and OC, becoming a crucial treatment outcome to assess [31]. Our systematic review identified seven HRQoL instruments developed or validated for OCC and OC patients. However, only five instruments achieved an overall score that met our threshold of 50.0 (EORTC QLQ-C30, EORTC QLQ-H&N43, EORTC QLQ-OH15, FACT-HN and UW-QoL).

Additionally, three instruments (i.e., EORTC QLQ-C30, EORTC QLQ-H&N43 and EORTC QLQ-OH15) received overall scores higher than 50.0 in all attributes evaluated. As a consequence, they are strongly recommended for use in the assessment of HRQoL in OCC and OC patients. The overall scores of these instruments were quite similar, with EORTC QLQ-H&N43 scoring slightly higher than the other two (84.1 vs. 80.0 and 73.7). The EORTC QLQ-H&N43 instrument is an updated version of EORTC QLQ-H&N35, which is widely used and has been translated and validated in several languages. It was updated because the developers detected some methodological issues, and some areas were improved in the updated version [32]. In the present study, the original instrument and the updated version were evaluated jointly, with an analysis of alternative modes of administration. Nevertheless, EORTC QLQ-H&N43 received the highest scores in all the domains, except the conceptual and measurement model, where it did not achieve the

highest score. However, it performed excellently, considering that the evaluated measurement concept and target population, as well as the dimensionality and distinction of the scales, were well described and argued.

EORTC QLQ-C30 showed good performance in all the domains evaluated; however, this instrument evaluates generic aspects of HRQoL for cancer patients, not specific for the head and neck [33], which can explain the worse internal consistency compared to EORTC QOL-C30 and EORTC QOL H&N43 [20]. Note that EORTC QOL-30 should be applied jointly with its head and neck module (EORTC H&N43) or its oral health module (EORTC QLQ-OH15) when evaluating HRQoL in patients with OCC and OC.

EORTC QLQ-OH15 presented a low score for responsiveness with 50.0. A lower score for responsiveness could be interpreted as a reduced ability of the instrument to detect a clinically significant change in HRQoL for the patient. EORTC QLQ-OH15 also exhibited lower interpretability than EORTC H&N43 and EORTC QOL-C30, primarily due to the absence of established minimally important difference (MID) points or guidelines for this module. It may be noted that there is no global rule for MID which is applicable to all environments, since MID patterns may differ due to the direction of change (improvement or deterioration) and the clinical environment [33]. However, for EORTC H&N43 and EORTC QOL-C30, MID guidelines are available only for patients with H&NC. There is a need to establish MID for OCC and OC patients.

FACT H&N consists of general questions about cancer with specific sub-scales for the head and neck. Its modular construction allows comparison between cancer diagnoses while investigating problems specific to oral cancer [34]. FACT H&N presented an overall score over 50.0, meaning that it is acceptable for evaluating HRQoL in patients with OCC and OC.

UW-QoL is specific for patients with H&NC, widely used and translated into several languages. However, we identified few studies that evaluated patients with OCC and OC. The lack of information from these studies affected the evaluation of some criteria, bringing its overall score down.

CnQ-SF-hn and QoL-OC received overall scores below 50.0. According to our results, none of them should be recommended for evaluating HRQoL in patients with OCC and OC. Future studies should concentrate on attributes like responsiveness, reliability, and validity to allow their re-evaluation for use in these patients. However, Gondivkar et al. [35] reported different results using the COSMIN (Consensus-based Standards for the selection of health Measurement Instruments) instrument. They concluded the QOL-OC was an integral PROM for measuring HRQoL in patients with OCC and that CnQ-SF-hn

Table 2 Rating of each EMPRO item and attribute for OHRQoL in Oral cavity cancer and Oropharyngeal cancer patients

	ATTRIBUTES	CnQ-SF-hn	EORTC QLQ-C30	EORTC QLQ-H&N43	EORTC QLQ-OH15	FACT H&N	QoL-OC	UW-QoL
	CONCEPT AND MEASURE- MENT MODEL	71.4	90.5	90.5	97.6	76.2	59.5	80.9
1	Concept of measurement stated	+++	++++	++++	++++	+++	+++	++++
2	Obtaining and combining items described	+++	++++	++++	++++	+++	+++	++
3	Rationale for dimensional- ity and scales	++	+++	+++	++++	+++	++	+++
4	Involvement of target population	++++	++++	++++	++++	+++	++	++++
5	Scale variability described and adequate	++	+++	+++	++++	+++	+++	+++
6	Level of measurement described	++	++++	++++	++++	+++	+++	++++
7	Procedures for deriving scores	++++	++++	++++	+++	++	++	++++
	RELIABILITY—total score	41.7	54.2	62.5	62.5	41.7	25.0	41.7
	Reliability: internal consistency	41.7	50.0	62.5	41.7	25.0	25.0	25.0
8	Data collection methods described	+++	++++	++++	++++	+++	++	++
9	Cronbach's alpha adequate	++++	++	+++	+++	++	++	+++
10	IRT estimates provided	n.a	n.a	n.a	n.a	n.a	n.a	n.a
11	Testing in different populations	-	+++	+++	-	-	-	-
	Reliability: reproducibility	29.2	54.2	58.3	62.5	41.7	20.8	41.7
12	Data collection methods described	+++	++++	+++	+++	++	++	++
13	Test–retest and time interval adequate	++	+++	++++	++++	++	+	+++
14	Reproducibility coeffi- cients adequate	+	++	+++	+++	+++	+	+++
15	IRT estimates provided	n.a	n.a	n.a	-	n.a	n.a	n.a
	VALIDITY	40.00	83.3	90.0	80.6	50.0	40.0	46.7
16	Content validity adequate	++++	+++	+++	++++	+++	++	++
17	Construct/criterion validity adequate	+++	+++	+++	+++	+++	+++	++
18	Sample composition described	++	++++	++++	++++	+++	+++	+++
19	Prior hypothesis stated	-	+++	++++	+++	+	++	+++
20	Rationale for criterion validity	n.a	n.a	n.a	++	n.a	n.a	n.a
21	Tested in different populations	-	++++	++++	++++	++	-	++
	RESPONSIVENESS	-	88.9	94.4	50.0	44.4	-	27.8
22	Adequacy of methods	-	++++	++++	+++	+++	-	++
23	Description of estimated magnitude of change	-	++++	++++	+++	++	-	++
24	Comparison of stable and unstable groups	-	+++	+++	+	++	-	-
	INTERPRETABILITY	22.2	83.3	83.3	77.8	61.1	22.2	55.6
25	Rationale of external criteria	++	+++	+++	++++	++	++	+++
26	Description of interpreta- tion strategies	+	+++	+++	++	+++	+	++

Table 2 (continued)

	ATTRIBUTES	CnQ-SF-hn	EORTC QLQ-C30	EORTC QLQ-H&N43	EORTC QLQ-OH15	FACT H&N	QoL-OC	UW-QoL
27	How data should be reported stated	++	+++	+++	++++	+++	++	+++
	OVERALL SCORE	35.6	80.0	84.1	73.7	54.7	29.3	50.5
	BURDEN							
	<i>Burden: respondent</i>	33.3	66.7	77.8	72.2	44.4	50.00	11.1
28	Skills and time needed	+++	+++	+++	+++	++	++	++
29	Impact on respondents	+	+++	+++	+++	++	++	+
30	Not suitable circumstances	+	+++	+++	+++	+++	+++	-
	<i>Burden: administrative</i>	75.0	100.0	100.0	75.0	91.7	37.5	33.3
31	Resources required	++	++++	++++	+++	+++	++	-
32	Time required	n.a	n.a	n.a	++++	n.a	++	n.a
33	Training and expertise needed	n.a	n.a	n.a	++	n.a	-	n.a
34	Burden of score calculation	++++	++++	++++	+++	++++	+++	+++
	ALTERNATIVE MODES OF ADMINISTRATION	-	-	100.00	100.00	-	-	-
35	Metric characteristics	n.a	n.a	++++	++++	n.a	n.a	n.a
36	Comparability	n.a	n.a	++++	n.a	n.a	n.a	n.a
	Scale of recommendation	Not recommended	Strongly recommended	Strongly recommended	Strongly recommended	Recommended with provisos or alterations	Not recommended	Recommended with provisos or alterations

Explanation: + + + + 4 (strongly agree); + + + 3; + + 2; + 1 (strongly disagree);—no information; n.a. not applicable. The higher the agreement the better the rating. Rows in white show EMPRO criteria assessing the results of the corresponding metric property, while rows in grey show EMPRO criteria assessing the methods applied to evaluate the corresponding metric property. Scale of recommendation following the criteria of Maratia et al. [30]: Strongly recommended (score higher than 50.00 in all the attributes), Recommended with provisos or alterations (overall EMPRO score higher than 50.00), and Not recommended (overall score lower than 50.00)

CnQ-SF-hn Cancer needs questionnaire, short form, head and neck cancer-specific, EORTC QLQ-C30 European Organisation for Research and Treatment of Cancer Quality of Life Core Questionnaire, EORTC QLQ-H&N43 European Organisation for Research and Treatment of Cancer head and neck cancer-specific quality of life questionnaire, EORTC QLQ-OH15 European organisation of Research and Treatment of Cancer, Oral Health Module, FACT-H&N (4.0) Functional assessment of cancer therapy scale, QOL-OC Oral cancer quality-of-life questionnaire, UW-QOL University of Washington quality of life questionnaire, OC Oral cancer

was suitable for evaluating patients with H&NC. In our evaluation, both instruments presented very low scores for reproducibility. Additionally, QOL-OC presented low scores for internal consistency and burden (administrative), which pulled down the overall score. The differences in the results between our investigation and the study of Gondivkar et al. [35] may be attributable to the tool used for the psychometric evaluation. Although both COSMIN and EMPRO focus on PRO, COSMIN evaluates the methodological quality of included studies rather than the instrument's measurement properties [36]. In contrast, EMPRO evaluates the instrument by assigning a numerical score reflecting the general quality of the measurement instrument [30].

Previous studies have identified that the worst scores are for interpretability, responsiveness [30, 37, 38], and burden [37]. However, our results indicated great weakness in responsiveness, mainly due to the lack of information on the magnitude of the change through comparative data with stable groups.

The only questionnaire we identified developed for patients with OCC (QOL-OC) obtained the worst overall score (29.3). Although the EMPRO evaluation indicated that this questionnaire needs to be improved, only one study was included in its assessment. Since the EMPRO score is influenced by the quantity and quality of the studies included, more studies evaluating metric properties are necessary before recommending its use.

The questionnaires that presented the best scores were developed for patients with H&NC. Although not designed for patients with OCC and OC, EORTC QLQ-H&N43, EORTC QLQ-C30, and EORTC QLQ-OH15 presented excellent psychometric properties when applied these patients. The good scores for the conceptual and measurement model, the good validity performance, the existence of MID criteria to facilitate interpretability, the ease of use, and good scores in responsiveness make these questionnaires recommendable for application in both cross-sectional and longitudinal studies.

Strengths and limitations

An important strength of this study is the evaluation of instruments using the EMPRO tool. EMPRO was designed to perform a general evaluation of the quality of the instruments, considering the methodology applied and the results obtained based on all the available evidence; it is a valid, reliable tool that has proven its usefulness for comparing the performance of both generics [26] and disease-specific PRO [26, 28, 30, 37, 38]. EMPRO has demonstrated strong internal consistency, agreement among raters, and positive associations that align with expected hypotheses between EMPRO attribute scores and quality indicators based on bibliometrics [26, 37]. Furthermore, this tool evaluates essential aspects of instrument use, such as reliability, validity and responsiveness [39].

It is important to consider these limitations when interpreting the findings of our study. The most obvious one is that it is possible that the search may not have identified all available health-related quality of life (HRQoL) instruments for patients with OCC and OC, due to language and database limitations. However, this limitation was addressed by searching specific databases PRO measures. Additionally, we applied a sensitive search strategy and carried out the review process in duplicate to minimize the impact of this limitation.

Another potential limitation was using a cut-off score to determine whether an instrument is appropriate. This classification was used to facilitate interpreting the results based on a study by Maratía et al. [30]: However, this approach can be subject to questioning. Furthermore, the assessors' personal experiences might impact the instruments' EMPRO assessment. Nevertheless, the assessments were conducted by investigators with wide experience in evaluating patient-reported outcome measures.

Lastly, instruments developed recently and with limited evidence due to their limited usage may have disadvantaged our evaluation. The quality and quantity of published evidence of included instrument significantly impact the EMPRO criteria assessment.

Conclusions

Based on the results of the standardised EMPRO evaluation, three instruments, namely the EORTC QLQ-H&N43, EORTC QLQ-C30 and EORTC QLQ-OH15, were found as appropriate for evaluating HRQoL in patients with OCC and OC in both cross-sectional and longitudinal studies. FACT-H&N and UW-QoL were found as appropriate for evaluating HRQoL in cross-sectional studies in these patients. Our results will provide clinicians and researchers with valuable guidance for selecting the most suitable instrument to evaluate HRQoL in patients with OCC and OC, thereby

contributing to a more accurate assessment of the impact of these cancers on patients' quality of life.

Abbreviations

CNQ-SF-hn (Chinese version)	Cancer needs questionnaire, short form, head and neck cancer-specific
COSMIM	COnsensus based Standards for the selection of health Measurement Instruments
EMPRO	Evaluating measures of patient-reported outcomes
EORTC QLQ-C30	European organization for research and treatment of cancer quality of life core questionnaire
EORTC QLQ-H&N43	European organization for research and treatment of cancer head and neck cancer specific quality of life questionnaire
EORTC QLQ-OH15	European organization of research and treatment of cancer, oral health module
FACT-H&N (4.0)	Functional assessment of cancer therapy scale;
QOL-OC	Oral cancer quality-of-life questionnaire
UW-QOL	University of Washington quality of life questionnaire
H&NC	Head and neck cancer
HRQoL	Health-related quality of life
OC	Oropharyngeal cancer
OCC	Oral cavity cancer
PRISMA	Preferred reporting items for systematic reviews and meta-analysis
HRQoL	Health-related quality of life

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12903-025-06984-2>.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements

ANID-Subdirección de Capital Humano/Doctorado Nacional/2021 [FOLIO21210983].

Authors' contributions

Conceptualization: N.F.D. and C.Z.; Methodology: N.F.D. and C.Z.; Formal analysis and investigation: N.F.D., C.U., F.C., N.A., R.M., W.G-A. and C.Z.; Writing – original draft preparation: N.F.D. and C.Z.; Writing—review and editing: N.F.D., C.U., F.C., N.A., R.M., W.G-A. and C.Z.; Funding acquisition: N.F.D. and C.Z.; Supervision: C.Z.

Funding

This project received financial support from the Universidad de La Frontera under grants PUA23-0006 and DI20-0054.

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author upon reasonable request.

Declarations

Ethics approval and consent to participate

This article does not contain any studies with human participants or animals performed by any of the authors.

Consent for publication

Not applicable.

Competing interests

Rodrigo Mariño is a Senior Editor for BMC Oral Health, and Carlos Zaror is an Editorial Board Member. The other authors declare no conflict of interest.

Author details

¹Center for Research in Epidemiology, Economics and Oral Public Health (CIEESPO), Faculty of Dentistry, Universidad de La Frontera, Temuco, Chile

²Department of Paediatric Dentistry and Orthodontics, Faculty of Dentistry, Universidad de La Frontera, Francisco Salazar Avenue #01145, Temuco 4811230, Chile

³Doctoral Program in Sciences in Applied Cellular and Molecular Biology, Universidad de La Frontera, Temuco, Chile

⁴Bucomaxillofacial Surgery and Traumatology Group, Complejo Asistencial Padre Las Casas, Padre Las Casas, Chile

⁵Faculty of Dentistry, Master Program in Dentistry, Universidad de La Frontera, Temuco, Chile

⁶Applied Morphology Research Centre (CIMA), Faculty of Dentistry, Universidad de La Frontera, Temuco, Chile

⁷Facultad de Medicina, Universidad de Atacama, Copiapó, Chile

⁸Center of Excellence in Surgical and Morphological Research (CEMyQ), Faculty of Medicine, Universidad de La Frontera, Temuco, Chile

⁹Department of Integral Adult Dentistry, Faculty of Dentistry, Universidad de La Frontera, Temuco, Chile

¹⁰Melbourne Dental School, University of Melbourne, Melbourne, Australia

¹¹Department of Conservative Dentistry and Oral Health, Riga Stradins University, Riga, Latvia

¹²Faculty of Dentistry, Universidad de Los Andes, Santiago, Chile

Received: 30 September 2024 / Accepted: 15 September 2025

Published online: 09 October 2025

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127(12):2893–917.
2. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–49. <https://doi.org/10.3322/caac.21660>.
3. Lin N-C, Hsien S-I, Hsu J-T, Chen MY. Impact on patients with oral squamous cell carcinoma in different anatomical subsites: a single-center study in Taiwan. *Sci Rep*. 2021;11(1):15446.
4. Alves N, Cândido P. Anatomia para o curso de Odontologia geral y específica. 4a ed. São Paulo: Gen-Santos; 2016. 360 p.
5. Gupta B, Bray F, Kumar N, Johnson NW. Associations between oral hygiene habits, diet, tobacco and alcohol and risk of oral cancer: a case-control study from India. *Cancer Epidemiol*. 2017;51:7–14. <https://doi.org/10.1016/j.canep.2017.09.003>.
6. Stepan KO, Mazul AL, Larson J, Shah P, Jackson RS, Pipkorn P, et al. Changing epidemiology of oral cavity cancer in the United States. *Otolaryngol Head Neck Surg*. 2023;168(4):761–8.
7. Creaney G, McMahon AD, Ross AJ, Bhatti LA, Paterson C, Conway DI. Head and neck cancer in the UK: what was the stage before COVID-19? UK cancer registries analysis (2011–2018). *Br Dent J*. 2022;233(9):787–93.
8. Gormley M, Creaney G, Schache A, Ingarfield K, Conway DI. Reviewing the epidemiology of head and neck cancer: definitions, trends and risk factors. *Br Dent J*. 2022;233(9):780–6.
9. National Cancer Institute; Surveillance, Epidemiology, and End Results Program. Cancer Stat Facts: Oral Cavity and Pharynx Cancer [Internet]. Bethesda (MD): National Cancer Institute. Available from: <https://seer.cancer.gov/statfacts/html/oralcav.html>. Cited 2025 Sep.
10. Glenny AM, Furness S, Worthington HV, Conway DI, Oliver R, Clarkson JE, et al. Interventions for the treatment of oral cavity and oropharyngeal cancer: radiotherapy. *Cochrane Database Syst Rev*. 2010;2010(12):CD006387. Epub 2010 Dec 8. <https://doi.org/10.1002/14651858.CD006387.pub2>. PMID: 21154367; PMCID: PMC10749265.
11. Parmar A, Macluskey M, McGoldrick N, Conway DI, Glenny AM, Clarkson JE, et al. Interventions for the treatment of oral cavity and oropharyngeal cancer: chemotherapy. *Cochrane Database Syst Rev*. 2021;12(12):CD006386. Epub 2021 Dec 20. <https://doi.org/10.1002/14651858.CD006386.pub4>. PMID: 34929047; PMCID: PMC8687638.
12. Worthington HV, Bulsara VM, Glenny AM, Clarkson JE, Conway DI, Macluskey M. Interventions for the treatment of oral cavity and oropharyngeal cancers: surgical treatment. *Cochrane Database Syst Rev*. 2023;8(8):CD006205. <https://doi.org/10.1002/14651858.CD006205.pub5>.
13. González-Arriagada WA, Ottaviani G, Dean D, Ottaviani G, Santos-Silva AR, Treister NS. Oral complications in cancer patients. *Front Oral Health*. 2023;3:116885. <https://doi.org/10.3389/froh.2022.116885>.
14. Epstein JB, Emerton S, Kolbinson DA, Le ND, Phillips N, Stevenson-Moore P, et al. Quality of life and oral function following radiotherapy for head and neck cancer. *Head Neck*. 1999;21(1):1–11.
15. Fingeret MC, Yuan Y, Urbauer D, Weston J, Nipomnick S, Weber R. The nature and extent of body image concerns among surgically treated patients with head and neck cancer. *Psychooncology*. 2012;21(8):836–44.
16. Moore K, Ford P, Farah C. Support needs and quality of life in oral cancer: a systematic review. *Int J Dent Hyg*. 2014;12(1):36–47.
17. Seedhouse D. Health: the foundations for achievement. Chichester (UK): John Wiley & Sons; 2001.
18. Gill TM, Feinstein AR. A critical appraisal of the quality of quality-of-life measurements. *JAMA*. 1994;272(8):619–26.
19. Yuwanati M, Gondivkar S, Sarode SC, Gadgil A, Desai A, Mhaske S, et al. Oral health-related quality of life in oral cancer patients: systematic review and meta-analysis. *Future Oncol*. 2021;17(8):979–90. <https://doi.org/10.2217/fon-2020-0881>.
20. Goiato MC, Amoroso AP, Silva B, dos Santos EG, de Caxias FP, Bitencourt SB, et al. The impact of surgery and radiotherapy on health-related quality of life of individuals with oral and oropharyngeal carcinoma and short-term follow up after treatment. *Asian Pac J Cancer Prev*. 2020;21(5):1227.
21. Kyrgidis A, Triaridis S, Kontos K, Patrikidou A, Andreadis C, Constantinidis J, et al. Quality of life in breast cancer patients with bisphosphonate-related osteonecrosis of the jaws and patients with head and neck cancer: a comparative study using the EORTC QLQ-C30 and QLQ-HN35 questionnaires. *Anticancer Res*. 2012;32(8):3527–34.
22. Rogers SN, Lowe D, Yueh B, Weymuller EA Jr. The physical function and social-emotional function subscales of the University of Washington Quality of Life Questionnaire. *Arch Otolaryngol Head Neck Surg*. 2010;136(4):352–7.
23. Ojo B, Genden EM, Teng MS, Milbury K, Misiukiewicz KJ, Badr H. A systematic review of head and neck cancer quality of life assessment instruments. *Oral Oncol*. 2012;48(10):923–37.
24. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Syst Rev*. 2021;10(1):1–11.
25. Liberati A, Altman D, Tetzlaff J, Mulrow C, Gøtzsche P, Ioannidis J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ*. 2009;339:b2700. <https://doi.org/10.1016/j.jclinepi.2009.06.006>.
26. Valderas JM, Ferrer M, Mendivil J, Garin O, Rajmil L, Herdman M, et al. Development of EMPRO: a tool for the standardized assessment of patient-reported outcome measures. *Value Health*. 2008;11(4):700–8. <https://doi.org/10.1111/j.1524-4733.2007.00309.x>.
27. Garin O, Herdman M, Vilagut G, Ferrer M, Ribera A, Rajmil L, et al. Assessing health-related quality of life in patients with heart failure: a systematic, standardized comparison of available measures. *Heart Fail Rev*. 2014;19(3):359–67. <https://doi.org/10.1007/s10741-013-9394-7>.
28. Schmidt S, Ferrer M, González M, González N, Valderas JM, Alonso J, et al. Evaluation of shoulder-specific patient-reported outcome measures: a systematic and standardized comparison of available evidence. *J Shoulder Elbow Surg*. 2014;23(3):434–44. <https://doi.org/10.1016/j.jse.2013.09.029>.
29. Schmidt S, Garin O, Pardo Y, Valderas JM, Alonso J, Rebollo P, et al. Assessing quality of life in patients with prostate cancer: a systematic and standardized comparison of available instruments. *Qual Life Res*. 2014;23(8):2169–81. <https://doi.org/10.1007/s11136-014-0678-8>.
30. Maratia S, Cedillo S, Rejas J. Assessing health-related quality of life in patients with breast cancer: a systematic and standardized comparison of available instruments using the EMPRO tool. *Qual Life Res*. 2016;25(10):2467–80.
31. van Nieuwenhuizen AJ, Buffart LM, Langendijk JA, Vergeer MR, Voortman J, Leemans CR, Verdonck-de Leeuw IM. Health-related quality of life and overall survival: a prospective study in patients with head and neck cancer

- treated with radiotherapy. *Qual Life Res.* 2021;30(4):1145–53. Epub 2020 Dec 8. <https://doi.org/10.1007/s11136-020-02716-x>. PMID: 33289866; PMCID: PMC8004520.
32. Singer S, Arraras JL, Chie W-C, Fisher SE, Galalae R, Hammerlid E, 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;22:1927–41.
 33. Musoro JZ, Coens C, Singer S, Tribius S, Oosting SF, Groenvold M, et al. Minimally important differences for interpreting European Organisation for Research and Treatment of Cancer quality of life questionnaire core 30 scores in patients with head and neck cancer. *Head Neck.* 2020;42(11):3141–52.
 34. Sheikh ME, Suleiman A, Satti A, O'Sullivan E. Translation, linguistic validation and reliability of FACT-H&N questionnaire in oral cancer patients in Sudan. *J Patient Rep Outcomes.* 2022;6(1):1–8.
 35. Gondivkar SM, Gadbail AR, Sarode SC, Gondivkar RS, Yuwanati M, Sarode GS, et al. Measurement properties of oral health-related patient-reported outcome measures in patients with oral cancer: a systematic review using COSMIN checklist. *PLoS ONE.* 2019;14(6):e0218833.
 36. Mokkink LB, Terwee CB, Gibbons E, Stratford PW, Alonso J, Patrick DL, et al. Inter-rater agreement and reliability of the COSMIN (Consensus-based Standards for the selection of health status Measurement Instruments) checklist. *BMC Med Res Methodol.* 2010;10:1–11.
 37. Deana NF, Pardo Y, Ferrer M, Espinoza-Espinoza G, Garin O, Muñoz-Millán P, et al. Evaluating conceptual model measurement and psychometric properties of oral health-related quality of life instruments available for older adults: a systematic review. *Health Qual Life Outcomes.* 2024;22(1):1–15.
 38. Zaror C, Pardo Y, Espinoza-Espinoza G, Pont Á, Muñoz-Millán P, Martínez-Zapata MJ, et al. Assessing oral health-related quality of life in children and adolescents: a systematic review and standardized comparison of available instruments. *Clin Oral Investig.* 2019;23(1):65–79. <https://doi.org/10.1007/s00784-018-2406-1>.
 39. Food and Drug Administration (FDA). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims; Silver Spring (MD): Food and Drug Administration, HHS. 2009.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.