

MR MOTAZ ASSAS (Orcid ID : 0000-0003-1206-8224)

DR PASWACH WIRIYAKIJJA (Orcid ID : 0000-0002-2340-517X)

DR STEFANO FEDELE (Orcid ID : 0000-0001-9006-9412)

Article type : Review Article

Title: Evaluating the measurement properties of patient-reported outcome measures in radiotherapy-induced xerostomia

Running Title: COSMIN systematic review of PROMs in RIX

Authors: Motaz Assas ^{1,2}, Paswach Wiriyakijja ^{2,3}, Stefano Fedele ^{2,4}, Stephen Porter ², Richeal Ní Ríordáin ²

Affiliations:

1. Department of Fundamental and Clinical Dental Science, Faculty of Dentistry, Umm Al Qura University, Makkah, Saudi Arabia.
2. Department of Oral Medicine, Eastman Dental Institute, UCL, London, United Kingdom.
3. Department of Oral Medicine, Faculty of Dentistry, Chulalongkorn University, Bangkok, Thailand.

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the [Version of Record](#). Please cite this article as [doi: 10.1111/ODI.13416](https://doi.org/10.1111/ODI.13416)

This article is protected by copyright. All rights reserved

4. National Institute for Health Research University College London Hospitals
Biomedical Research Centre

Corresponding author:

Motaz Assas (Motaz.Assas.13@ucl.ac.uk)

Oral Medicine Unit

UCL Eastman Dental Institute

256 Gray's Inn Road,

London WC1X 8LD, UK

Tel: +44(0)2034567890 (Ext: 61193)

Keywords: Validity, Quality of life, Xerostomia, Radiotherapy, COSMIN.

Date of Submission: 07/01/2020

Abstract

Objective: Radiotherapy-induced xerostomia (RIX) is one of the most common adverse effects of radiotherapy to the head and neck, and a major determinant of survivors' quality of life. A number of patient-reported outcome measures (PROMs) have been used in clinical trials of therapeutic interventions for RIX; however, little is known regarding their measurement properties and methodological quality. *Methods:* We conducted a systematic literature search in Embase, Medline and PsycINFO for articles published up to May 2019 and evaluating at least one measurement property of PROMs relevant to RIX. The COSMIN guidelines were used to assess relevant measurement properties and methodological quality. *Results:* Nine validation studies were identified reporting on four PROMs relevant to RIX. The Xerostomia Questionnaire (XQ) showed overall high-quality evidence for structural

validity and internal consistency, but low-quality evidence supporting reliability. The methodological quality of the Groningen Radiotherapy-Induced Xerostomia scale (GRIX), Xerostomia Inventory (XI) and the Xerostomia Quality of life scale (XeQoLS) was of relatively low-quality for all measurement properties. *Conclusions:* The XQ was found to have the highest potential to capture changes in RIX according to COSMIN guidelines. Additional validation studies are required to further understand the methodological quality of the XI, GRIX and XeQoLS.

Introduction

Xerostomia is a common permanent adverse effect of radiotherapy to the head and neck. Radiotherapy-induced xerostomia (RIX) in head and neck cancer (HNC) survivors can affect speech and eating, cause persistent discomfort, and increase the risk of infections and dental disease, with consequent negative impact upon the quality of life (QoL) of affected individuals. (Fang, Liu, Tang, Wang, & Ko, 2004; Jensen, Bonde Jensen, & Grau, 2006). Measurement of RIX includes the subjective assessment of the severity of dry mouth symptoms, as well as the QoL of affected individuals, via patient-reported outcome measures (PROMs) (Ringash et al., 2015). In order for clinicians to be reassured that a PROM can adequately measure the symptom of interest, validation studies should be performed and its measurement properties should be assessed (Mokkink et al., 2018). According to the Consensus-based standards for the selection of health measurement Instruments (COSMIN), assessment of the measurement properties of PROMs should include their reliability, validity and responsiveness (Mokkink et al., 2006; Mokkink et al., 2010; Prinsen et al., 2018). Little is known regarding the measurement properties of PROMs relevant to RIX. In 2012 Ojo et al used the Scientific Advisory Committee of the Medical Outcome Trust (SAC-MOT) guidelines in order to assess the properties of QoL instruments in head and neck cancer, some of which included the RIX dimension (Ojo et al., 2012). Their

results showed a lack of rigorous testing for the instruments measuring RIX (Caroline B. Terwee et al., 2011).

In this review, we use the COSMIN guidelines so to critically appraise, pool and compare the measurement properties of all available PROMs measuring RIX (Prinsen et al., 2018). The aim is to provide clinicians with an evidence-based recommendation regarding the most suitable PROMs measuring RIX outcomes in clinical practice and future research.

Methods

Search strategy

A systematic literature search was performed on Embase, Medline and PsycINFO up until May 2019. The search strategy consisted of three search filters:

1. Construct: Radiotherapy-induced xerostomia.
2. Target population: HNC patients treated with radiotherapy
3. A validated search filter for instrument measurement properties (validity, reliability, and responsiveness) in Embase, Medline and PsycINFO (C. B. Terwee, Jansma, Riphagen, & de Vet, 2009).

Inclusion and exclusion

All titles and abstracts from the search result were screened by one of the authors (MA) and then (MA and PW) assessed the full texts for eligibility. Articles were included based on the following criteria:

1. Construct: The PROM aim is to measure any aspects of xerostomia in HNC patients based upon their perspective. Any instruments measuring other symptoms or using a non-xerostomia specific PROM were excluded.
2. Target population: HNC patients that have developed RIX exclusively by receiving radiotherapy as a module of treatment. Any other xerostomia patients (e.g. Sjogren syndrome or medication-induced salivary gland dysfunction) were excluded.
3. Study aim: Articles related to the evaluation of 1 or more measurement properties (validity, reliability, and responsiveness) and its development were included while studies

with insufficient methods of validity or using methods unidentifiable based on the COSMIN guidelines were excluded (Mokkink et al., 2010).

Assessment of methodological quality of included studies

First, the COSMIN taxonomy was used to establish which measurement properties were assessed in each included study based on its definition (Table 1). The characteristics of each included PROM and characteristics of the included study populations were extracted. The COSMIN risk of bias was then used to assess the methodological quality of the included studies (Mokkink et al., 2018), scoring each measurement property rated as very good, adequate, doubtful or inadequate. MA and PW performed the evaluation, and a third reviewer (RNR) was reached in case of disagreement.

Assessment of measurement property results

The scores from the methodological quality assessment are then compared against the criteria of good measurement quality (Mokkink et al., 2018).

Content validity is thought to be the most critical measurement property since the items in a PROM should first be clear, relevant, comprehensive and easily understandable before ensuring an appropriate internal structure (Mokkink et al., 2018; C. B. Terwee et al., 2018). With the absence of this information, we, therefore, did not report on content validity. To help with interpreting hypotheses for hypothesis testing, responsiveness and assessing correlations, these pre-defined hypotheses were set:

1. Correlations measuring a PROM against a similar construct should at least be greater than 0.50.
2. There should be a significantly reported change in correlations between subgroup and changes over time.

Evidence synthesis

The summarised results are then evaluated against the criteria for good measurement properties to get an overall rating; sufficient (+) insufficient (–) and indeterminate or Inconsistent (?) for each the measurement property. Finally, the GRADE approach (Grading

of Recommendations Assessment, Development and Evaluation) is applied (Prinsen et al., 2018). The GRADE level of evidence (High, Moderate, Low and Very low) is based on four points:

- Risk of bias: as determined using the COSMIN Risk of Bias checklist (the more the risk, the less trustworthiness of the evidence).
- Inconsistency of pooled results.
- Imprecision: assessing if the sample size is large or small ($n > 100$).
- Indirectness: this refers to evidence coming from different populations.

Results

One hundred seventy-eight articles were included in the search strategy. 148 articles were screened (30 duplicates) and 142 articles were excluded because of the following reasons; not assessing measurement properties ($n=78$), using a non-specific xerostomia PROM instrument ($n=33$), review studies ($n=18$), not measuring xerostomia ($n=13$). From reference checking, three articles that met the inclusion criteria were added (indicated in Figure 1).

A total of 9 articles with four different PROMs were assessed and scored in this review. Two PROMs, the Xerostomia Questionnaire (XQ) and the Xerostomia Inventory (XI), measured the severity of xerostomia, whereas the Groningen Radiotherapy-induced Xerostomia Questionnaire (GRIX) and the Xerostomia Quality of Life Scale (XeQoLS) focussed on the impact of xerostomia on patient's QoL. Characteristics of the included PROMs are shown in Table 2. Characteristics of the study populations are detailed in Table 3. The quality of evidence for each measurement properties result for each PROM is presented in Table 4. Instruments were placed in order of construct, severity and QoL outcome. A supplementary electronic table is available in the online version of this review (Prinsen et al., 2018), it includes the results of all the included articles on measurement properties for each PROM.

All included PROMs were found to have evidence supporting measurement properties including internal consistency, reliability, construct validity and responsiveness, except for the XeQoLS, which did not have reports on reliability. XQ was the only PROM found to have evidence supporting its structural validity. The included studies in this review did not explore content validity (C. B. Terwee et al., 2018). Additionally, with lack of a standardised consensus on what represents content validity in xerostomia research, we could not report on content validity in this review. The summary of each PROM with evidence supporting their measurement properties for use in HNC patients with RIX is outlined below.

Groningen Radiotherapy-Induced Xerostomia (GRIX)

The GRIX questionnaire assesses different aspects of patient-reported xerostomia to evaluate the impact of radiotherapy on patients' QoL due to the emergence of new modes of radiation delivery (Beetz et al., 2010). GRIX contains 14 items with four subscales, 2 for xerostomia and 2 for sticky saliva, which is measured twice for each during daytime and at night. A 4-point Likert-like scale (not at all, a little, quite a bit, very much) is used to score each item. Scores are then converted to a 0-100 scale, with higher scores indicating a greater degree of patient-reported xerostomia.

The GRIX was found to have an indeterminate overall scoring of low quality evidence for internal consistency and construct validity, an indeterminate overall scoring of very low quality evidence for reliability and a sufficient overall scoring of very low quality evidence for responsiveness.

Xerostomia Questionnaire (XQ)

The XQ was designed to measure xerostomia by patients rating their symptom severity. It consists of 8 items, four items about dryness while eating and chewing, and the other four about dryness while not eating or chewing. Each item is rated on an 11-point numerical rating scale from 0 to 10, with higher scores indicating more dryness or more dryness discomfort. A final summary score is derived from summing the item scores and transforming it linearly to a 0-100 score, and the greater the score, the higher the xerostomia severity (Eisbruch et al., 2001).

The XQ was found to have a sufficient overall scoring of high quality evidence for structural validity and internal consistency. Reliability of the instrument was found to have low quality evidence with an indeterminate overall scoring. A sufficient overall scoring was found for construct validity; with moderate quality evidence, and responsiveness; with low quality evidence.

Xerostomia Inventory (XI)

The Xerostomia Inventory is designed to measure the severity of symptoms associated with xerostomia. XI contains 11 items with a 5-point Likert-like response scale of never (scoring 1), hardly ever (2), occasionally (3), fairly often (4) and very often (5) (W. M. Thomson, 2007; Thomson & Williams, 2000).

A sufficient overall scoring for XI was found for internal consistency; with moderate quality evidence, and construct validity; with very low quality evidence, and responsiveness; with low quality evidence. Reliability was found to have an indeterminate overall scoring with very low quality evidence.

Xerostomia Quality of Life Scale (XeQoLS)

The Xerostomia Quality of Life Scale is developed to measure QoL in patients with xerostomia. It consists of 15 items with a 5 Likert-like response scales (not at all, a little, somewhat, quite a bit, very much). This tool represents 4 QoL dimensions; physical functioning, pain/discomfort, personal/psychological functioning and social functioning (B. S. I. M. R. E. Henson, 2001). The XeQoLS was found to have indeterminate overall score with very low quality evidence for internal consistency, and a sufficient overall scoring with very low quality evidence for construct validity and responsiveness.

Discussion

The COSMIN has recently published new guidelines for systemically reviewing PROMs and reporting on measurement properties and its evidential rigour (Mokkink et al., 2018; Prinsen et al., 2018). The application of the COSMIN guidelines can facilitate the selection of PROMs to be used in clinical research and can highlight areas of further psychometric development that might be required.

The present review has identified validation studies relevant to a total of four PROMs measuring RIX in HNC patients: the XQ, XI, GRIX and XeQoLS. These four PROMs were developed so to measure various aspects of xerostomia. For example, the XQ focuses on the intensity of xerostomia during eating and chewing (Eisbruch et al., 2001) whereas the XI measures a range of symptoms related to xerostomia such as itching, dryness, and burning sensation (Thomson & Williams, 2000). The XeQoLS aims to capture the impact of xerostomia on the QoL of patients, with domains including physical functioning, pain and psychosocial functioning (Duke et al., 2004). GRIX is the only tool explicitly developed in the RIX population, and focuses on temporal aspects (daytime vs night-time) of xerostomia and sticky saliva (Beetz et al., 2010). Of note, XQ is the most frequently used questionnaire in studies measuring RIX (Eisbruch et al., 2001; Eisbruch et al., 2003; Hawkins et al., 2018; Kamal et al., 2018; A. Lin et al., 2003; S.-C. Lin, Jen, Chang, & Lin, 2008; Meirovitz, Murdoch-Kinch, Schipper, Pan, & Eisbruch, 2006; Memtsa et al., 2017; Pellegrino, Groff, Bastiani, Fattori, & Sotti, 2015; Trotti & Eisbruch, 2011). Critical assessment of the measurement properties of the above PROMs has presented a number of challenges. The XQ has been difficult to assess as the relevant papers have described it as having “an 11-point ordinal Likert scale” from 0-10 with a two point threshold at start and finish (e.g. Easy to Extremely difficult or No dryness to Extreme dryness) (Eisbruch et al., 2001). However a Likert-like scale is described as a scale with meanings or descriptions attached to each point on the scale, and not a numerical scale with a description on the start and end point (Allen & Seaman, 2007). Therefore, we decided to consider the XQ as having a numerical scale and assessed it based on this description. Furthermore, the evaluation of criterion validity for the GRIX validation study was considered by the authors of the present review as an evaluation of construct validity instead. Criterion validity is described by the COSMIN taxonomy as “the degree to which the scores of a PROM are an adequate reflection of a ‘gold standard’” (Mokkink et al., 2010). With the absence of comparison between the GRIX and its golden standard, the COSMIN guideline recommends evaluating it instead as hypothesis testing for construct validity.

The results of our review suggest that the XQ is the PROM of RIX best suited for further use in clinical interventions, as it has high level of evidence for structural validity and internal consistency, as well as a sufficient overall score with moderate quality of evidence for responsiveness. However, XQ was found to have an indeterminate overall score; of low quality evidence, for reliability. Therefore future well-designed research examining the reliability of XQ is needed. The GRIX, XI and XeQoLS instruments are in need of further validation studies as they were found to have low scores of both methodological quality and quality rating.

The present study has a number of limitations. The COSMIN guidelines provide evidence-based recommendations so to identify PROMs that have been tested through rigorous methodology. However, it should be noted that not reporting on some measurement properties, or the low scoring of specific measurement properties of a PROM does not necessarily mean that the instrument has been poorly developed; instead it might indicate that the relevant validation study was poorly-designed and executed. This could be improved by further testing of the measurement properties of the instrument, without the need of re-development.

Another limitation of this review is the absence of content validity. The COSMIN guidelines encourage reviewers to seek content validity, even when this information is missing in the relevant validation study. This can be achieved by asking patients and clinicians on the relevancy and comprehensiveness of a given PROM, and also asking patients on its comprehensibility. Relevant information can then be assessed separately using the COSMIN standard for evaluating content validity (C. B. Terwee et al., 2018). This evaluation, which was not part of the present review, could be performed by future studies with a view to complete the assessment of the methodological quality of the above PROMs.

Conclusion

The XQ was found to be the PROM with the highest potential to capture changes in RIX according to COSMIN guidelines, although there is a need for future research to assess its content validity. Further validation studies are required in order to better understand the methodological quality of the GRIX, XI and XeQoLS.

Conflict of interest

Stefano Fedele received funding from the National Institute for Health Research University College London Hospitals Biomedical Research Centre. Richeal Ní Ríordáin received funding from the National Institute for Health Research.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.



Figure 1 PRISMA 2009 Flow Diagram

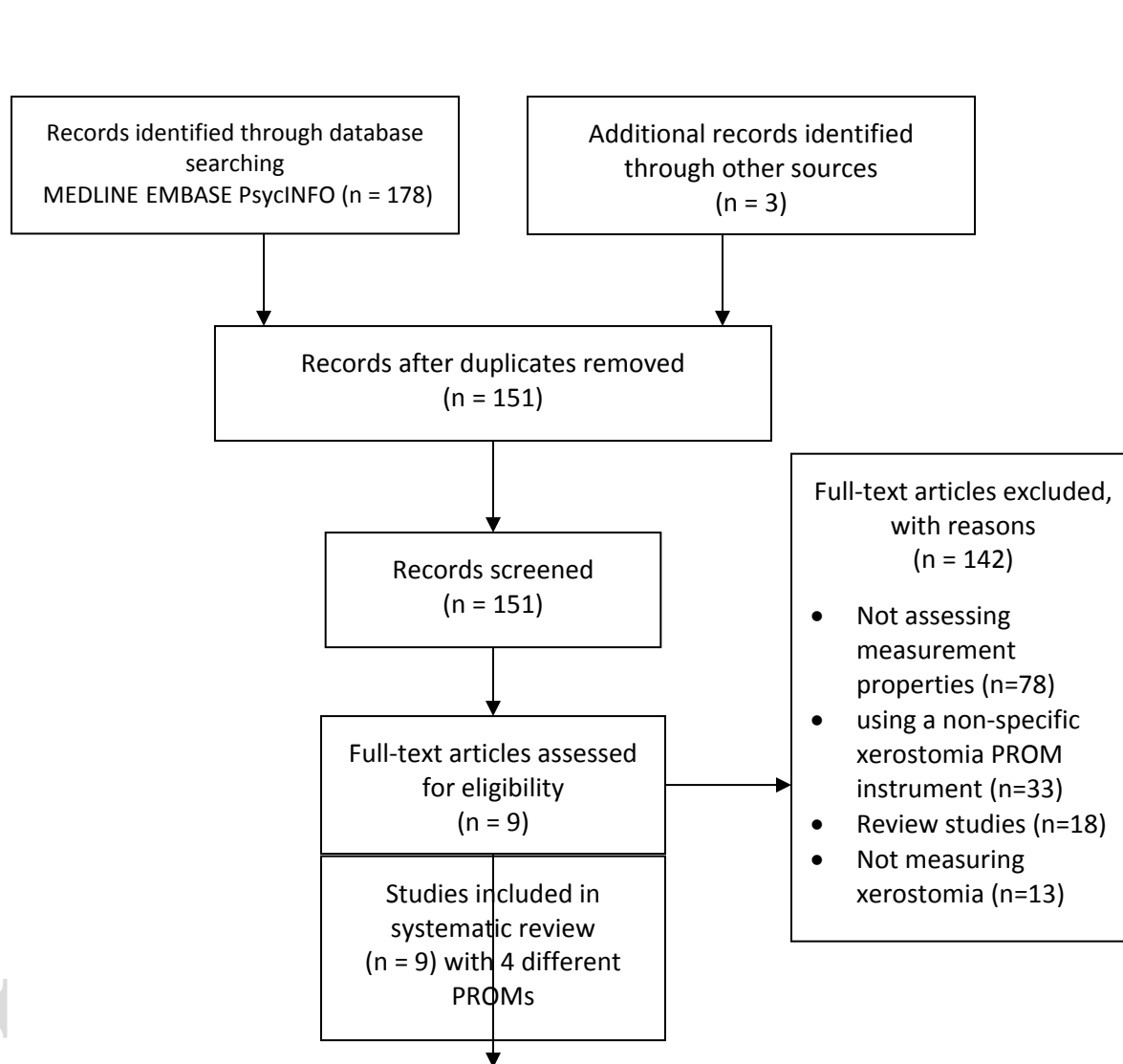


Table 1. COSMIN definitions of domains, measurement properties, and aspects of measurement properties (Mokkink et al., 2010)

Term			Definition
Domain	Measurement property	Aspects of measurement property	
Reliability			The degree to which the measurement is free from measurement error
Reliability (extended definition)			The extent to which scores for patients who have not changed are the same for repeated measurement under several conditions: e.g. using different sets of items from the same PROM (internal consistency); over time (test-retest); by different persons on the same occasion (interrater); or by the same persons (i.e. raters or responders) on different occasions (intra-rater)
	Internal consistency		The degree of the interrelatedness among the items
	Reliability		The proportion of the total variance in the measurements which is due to 'true' differences

			between patients
	Measurement error		The systematic and random error of a patient's score that is not attributed to true changes in the construct to be measured
Validity			The degree to which a PROM measures the construct(s) it purports to measure
	Content validity		The degree to which the content of a PROM is an adequate reflection of the construct to be measured
		Face validity	The degree to which (the items of) a PROM indeed looks as though they are an adequate reflection of the construct to be measured
	Construct validity		The degree to which the scores of a PROM are consistent with hypotheses (<i>for instance with regard to internal relationships, relationships to scores of other instruments, or differences</i>

			<i>between relevant groups</i>) based on the assumption that the PROM validly measures the construct to be measured
		Structural validity	The degree to which the scores of a PROM are an adequate reflection of the dimensionality of the construct to be measured
		Hypotheses testing	Idem construct validity
		Cross-cultural validity	The degree to which the performance of the items on a translated or culturally adapted PROM are an adequate reflection of the performance of the items of the original version of the PROM
	Criterion validity		The degree to which the scores of a PROM are an adequate reflection of a 'gold standard'
Responsiveness			The ability of a PROM to detect change over time in the construct to be measured
	Responsiveness		Idem responsiveness
Interpretability			Interpretability is the degree to which one can assign qualitative meaning - that is, clinical or commonly understood connotations – to a PROM's quantitative scores or change in scores.

PROM	Construct(s)	Target population	conceptual model used	Recall period	(Sub)scale (s) (number of items)	Response options	Range of scores/scoring	Available translations
GRIX (Beetz et al., 2010)	QoL related to xerostomia	HNC patients with RIX	Existing PROMs reviewed, expert opinion and a single patient opinion	Not specified	Four subscales: Xerostomia during the night, Xerostomia during the daytime, Sticky saliva during the night and Sticky saliva during daytime (14)	4 – point Likert, 0-3	All scores converted linearly to a 0-100 scale, with higher scores representing more xerostomia	English
XQ (Eisbruch et al., 2001)	Severity of xerostomia	Patients with xerostomia	Existing xerostomia-specific and general HNC QoL instruments PROMs reviewed surveys of patients and discussions with members of the research team.	Seven days	Two subscales: Dryness while eating or chewing and Dryness while not eating and chewing (8)	11-point numerical rating scale, 0-10	All scores converted linearly to a 0-100 scale, with higher scores representing more xerostomia	Italian, Taiwanese, Persian, Greek, French
XI (Thomson & Williams, 2000)	Severity of xerostomia	Patients with xerostomia	Used in HNC patients based on the validation study of the instrument on other dry mouth patients.	Two weeks	Symptom frequency (11)	5-point Likert, 1-5	Higher scores indicate severe xerostomia	Non-available

XeQoLS (B. S. Henson, Inglehart, Eisbruch, & Ship, 2001)	QoL related to xerostomia	Patients with xerostomia	Based on previous studies on QoL	Not specified	Physical functioning, Pain/discomfort issues, Personal/psychological functioning, Social functioning (15)	Not at all, a little, somewhat, quite a bit, very much	Averaging the values of all the respective items for that individual domain, a total average calculated	Italian
--	---------------------------	--------------------------	----------------------------------	---------------	---	--	---	---------

Table 2 characteristics of the included PROM¹

¹ GRIX, Groningen Radiotherapy-Induced Xerostomia; NA no available data; XeQoLS, Xerostomia Quality of Life Scale; XI, Xerostomia Inventory; XQ, Xerostomia Questionnaire.

Table 3 characteristics of the included study population²

PROM	Reference	Population			PROM administration			Response rate
		N	Age Mean, (SD and/or range in years)	Gender % female	Setting	Country	Language	
GRIX	(Beetz et al., 2010)	315	62 years, (19-90 years)	31%	Department of radiation oncology of the university medical centre of Groningen	The Netherlands	Dutch	Not reported
XQ	(Eisbruch et al., 2001)	132	51 years	31%	The University of Michigan	USA	English	Not reported
XQ-T	(S.-C. Lin et al., 2008)	50	54 years (SD 14.42)	16%	The radiology oncology outpatient clinic of a medical centre	Taipei, Taiwan	Taiwanese	Not reported
XQ-G	(Memtsa et al., 2017)	100	63.4 years (SD 7.5)	27%	Radiation Therapy Departments of University Hospitals of Larissa, Theagenio hospital of Thessaloniki and AXEPA hospital of Thessaloniki	Greece	Greek	Not reported
XQ-IT	(Pellegrino et al., 2015)	102	62.9 years (24–85 years)	18.6%	Radiotherapy Unit of the Veneto Oncology Institute-IOV	Padua, Italy	Italian	Not reported
XI	(Thomson & Williams, 2000)	112	<ul style="list-style-type: none"> Onset group 63 years (SD 13; 29-87 years) Normal group 75 (SD 7; 52-90 years). 	Onset group 28.1% Normal group 32.7%	Radiotherapy units at each of Auckland, Waikato, Palmerston North, Wellington, Christchurch, and Dunedin hospitals. Controls from the membership list of the Otago Medical Research Foundation Auxiliary	New Zealand	English	72.2%

² XQ-T, XQ Taiwan version; XQ-G, Greek version; XQ-IT, XQ Italian version; XeqoLS-IT, XeqoLS Italian version

XI	(W. Murray Thomson, 2007)	94	68.6 years (SD, 12.9; R 29–90 years)	Onset group 28.3% Normal group 38.1%	Radiotherapy units at each of Auckland, Waikato, Palmerston North, Wellington, Christchurch, and Dunedin hospitals. Controls from the membership list of the Otago Medical Research Foundation Auxiliary	New Zealand	English	Not reported
XeQoLS	(B. S. Henson et al., 2001)	20	55.8 years (SD 12.8; 24–80 years)	42%	the University of Michigan Radiation–Oncology	USA	English	Not reported
XeQoLS-IT	(Lastrucci et al., 2018)	35	R 18–83 years	17.2%	Unit of Radiation Oncology, S. Donato Hospital	Arezzo, Italy	Italian	Not reported

Table 4 Summary of findings and overall evidence score^{3,4}

PROM	Structural Validity	Internal Consistency	Measurement Invariance	Reliability	Measurement Error	Validity		Responsiveness
						Criterion	Construct	
Severity Outcome								
XQ	+ (High)	+ (High)	NR	? (Low)	NR	NR	+ (Moderate)	+ (Low)

³ Level of evidence, High indicates that we are very confident that the true measurement property lies close to that of the estimate/pooled result of measurement property; moderate, we are moderately confident in the measurement property estimate: the true measurement property is likely to be close to the estimate of the measurement property, but there is a possibility that it is substantially different; low, our confidence in the measurement property estimate is limited: the true measurement property is likely to be substantially different from the estimate of the measurement property; very low. We have little confidence in the measurement property estimate: the true measurement property is likely to be substantially different from the estimate of the measurement property.

⁴(+) Sufficient overall measurement property rating; (?) indeterminate overall measurement property rating; (NR) Not reported.

XI	NR	+ (Moderate)	NR	? (Very low)	NR	NR	+ (Low)	+ (Low)
QoL Outcome								
GRIX	NR	? (Low)	NR	? (Very low)	NR	NR	? (Low)	+ (Very low)
XeQoLS	NR	? (Very low)	NR	NR	NR	NR	+ (Very low)	+ (Very low)

References

- Allen, I., & Seaman, C. (2007). Likert scales and data analyses. *Quality progress*, 40(7), 64-65.
- Beetz, I., Burlage, F., Bijl, H., Hoegen-Chouvalova, O., Christianen, M., Vissink, A., . . . Langendijk, J. (2010). The Groningen Radiotherapy-Induced Xerostomia questionnaire: Development and validation of a new questionnaire. *Radiotherapy and Oncology*, 97(1), 127-131. doi:http://dx.doi.org/10.1016/j.radonc.2010.05.004
- Duke, R. L., Campbell, B. H., Schultz, C. J., Marbella, A., Myers, K. B., & Layde, P. (2004). Xerostomia and effects on quality of life in long-term head and neck cancer survivors. *Otolaryngology - Head and Neck Surgery*, 131(2), P183-P184. doi:http://dx.doi.org/10.1016/j.otohns.2004.06.334
- Eisbruch, A., Kim, H. M., Terrell, J. E., Marsh, L. H., Dawson, L. A., & Ship, J. A. (2001). Xerostomia and its predictors following parotid-sparing irradiation of head-and-neck cancer. *International Journal of Radiation Oncology*Biology*Physics*, 50(3), 695-704. doi:http://dx.doi.org/10.1016/S0360-3016(01)01512-7
- Eisbruch, A., Rhodus, N., Rosenthal, D., Murphy, B., Rasch, C., Sonis, S., . . . Brizel, D. (2003). The prevention and treatment of radiotherapy-induced xerostomia. *Seminars in Radiation Oncology*, 13(3), 302-308. doi:http://dx.doi.org/10.1016/S1053-4296(03)00027-4
- Fang, F. M., Liu, Y. T., Tang, Y., Wang, C. J., & Ko, S. F. (2004). Quality of life as a survival predictor for patients with advanced head and neck carcinoma treated with radiotherapy. *Cancer*, 100(2), 425-432. doi:10.1002/cncr.20010
- Hawkins, P. G., Lee, J. Y., Mao, Y., Li, P., Worden, F., Swiecicki, P. L., . . . Eisbruch, A. (2018). Sparing All Salivary Glands With IMRT for Head and Neck Cancer: Longitudinal Study of Patient-Reported Xerostomia and Head and Neck Quality Of Life. *International Journal of Radiation Oncology • Biology • Physics*, 100(5), 1388-1389. doi:10.1016/j.ijrobp.2017.12.210

- Henson, B. S., Inglehart, M. R., Eisbruch, A., & Ship, J. A. (2001). Preserved salivary output and xerostomia-related quality of life in head and neck cancer patients receiving parotid-sparing radiotherapy. *Oral Oncology*, 37(1), 84-93.
- Henson, B. S. I. M. R. E. (2001). Preserved salivary output and xerostomia-related quality of life in head and neck cancer patients receiving parotid-sparing radiotherapy. *Oral Oncology*, 37(1), 84-93.
- Jensen, K., Bonde Jensen, A., & Grau, C. (2006). The relationship between observer-based toxicity scoring and patient assessed symptom severity after treatment for head and neck cancer. A correlative cross sectional study of the DAHANCA toxicity scoring system and the EORTC quality of life questionnaires. *Radiother Oncol*, 78(3), 298-305. doi:10.1016/j.radonc.2006.02.005
- Kamal, M., Rosenthal, D., Volpe, S., Goepfert, R., Garden, A., Hutcheson, K., . . . Gunn, G. (2018). Patient reported dry mouth: Instrument comparison and model performance for correlation with quality of life in head and neck cancer survivors. *Radiotherapy and Oncology*, 126(1), 75-80. doi:https://doi.org/10.1016/j.radonc.2017.10.037
- Lastrucci, L., Bertocci, S., Bini, V., Borghesi, S., De Majo, R., Rampini, . . . Pernici, P. (2018). Xerostomia Quality of Life Scale (XeQoLS) questionnaire: validation of Italian version in head and neck cancer patients. *Radiol Med*, 123(1), 44-47. doi:10.1007/s11547-017-0798-7
- Lin, A., Kim, H. M., Terrell, J. E., Dawson, L. A., Ship, J. A., & Eisbruch, A. (2003). Quality of life after parotid-sparing IMRT for head-and-neck cancer: A prospective longitudinal study. *International Journal of Radiation Oncology*Biology*Physics*, 57(1), 61-70. doi:https://doi.org/10.1016/S0360-3016(03)00361-4
- Lin, S.-C., Jen, Y.-M., Chang, Y.-C., & Lin, C.-C. (2008). Assessment of Xerostomia and Its Impact on Quality of Life in Head and Neck Cancer Patients Undergoing Radiotherapy, and Validation of the Taiwanese Version of the Xerostomia Questionnaire. *Journal of Pain and Symptom Management*, 36(2), 141-148. doi:https://doi.org/10.1016/j.jpainsymman.2007.09.009

- Meirovitz, A., Murdoch-Kinch, C. A., Schipper, M., Pan, C., & Eisbruch, A. (2006). Grading xerostomia by physicians or by patients after intensity-modulated radiotherapy of head-and-neck cancer. *International Journal of Radiation Oncology*Biology*Physics*, 66(2), 445-453. doi:http://dx.doi.org/10.1016/j.ijrobp.2006.05.002
- Memtsa, P. T., Tolia, M., Tzitzikas, I., Bizakis, I., Pistevou-Gombaki, K., Charalambidou, M., . . . Kyrgias, G. (2017). Validity and reliability of the Greek version of the xerostomia questionnaire in head and neck cancer patients. *Supportive Care in Cancer*, 25(3), 847-853. doi:10.1007/s00520-016-3471-0
- Mokkink, L., De Vet, H., Prinsen, C., Patrick, D., Alonso, J., Bouter, L., & Terwee, C. (2018). COSMIN Risk of Bias checklist for systematic reviews of Patient-Reported Outcome Measures. *Qual Life Res*, 27(5), 1171-1179. doi:10.1007/s11136-017-1765-4
- Mokkink, L., Terwee, C., Knol, D., Stratford, P., Alonso, J., Patrick, D., . . . De Vet, H. (2006). Protocol of the COSMIN study: COnsensus-based Standards for the selection of health Measurement INstruments. *BMC Medical Research Methodology*, 6, 2-2. doi:10.1186/1471-2288-6-2
- Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., . . . De Vet, H. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*, 63(7), 737-745. doi:10.1016/j.jclinepi.2010.02.006
- Ojo, B., Genden, E. M., Teng, M. S., Milbury, K., Misiukiewicz, K. J., & Badr, H. (2012). A systematic review of head and neck cancer quality of life assessment instruments. *Oral Oncology*, 48(10), 923-937.
- Pellegrino, F., Groff, E., Bastiani, L., Fattori, B., & Sotti, G. (2015). Assessment of radiation-induced xerostomia: validation of the Italian version of the xerostomia questionnaire in head and neck cancer patients. *Supportive Care in Cancer*, 23(4), 925-932. doi:10.1007/s00520-014-2438-2
- Prinsen, C. A. C., Mokkink, L. B., Bouter, L. M., Alonso, J., Patrick, D. L., de Vet, H. C. W., & Terwee, C. B. (2018). COSMIN guideline for systematic reviews of patient-reported

outcome measures. *Qual Life Res*, 27(5), 1147-1157. doi:10.1007/s11136-018-1798-3

Ringash, J., Bernstein, L. J., Cella, D., Logemann, J., Movsas, B., Murphy, B., . . . Ridge, J. (2015). Outcomes toolbox for head and neck cancer research. *Head and Neck*, 37(3), 425-439.

Terwee, C. B., Jansma, E. P., Riphagen, II, & de Vet, H. C. (2009). Development of a methodological PubMed search filter for finding studies on measurement properties of measurement instruments. *Qual Life Res*, 18(8), 1115-1123. doi:10.1007/s11136-009-9528-5

Terwee, C. B., Mokkink, L. B., Knol, D. L., Ostelo, R. W. J. G., Bouter, L. M., & Vet, H. C. W. (2011). Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. *Quality of Life Research*, 21(4), 651-657. doi:10.1007/s11136-011-9960-1

Terwee, C. B., Prinsen, C. A. C., Chiarotto, A., Westerman, M. J., Patrick, D. L., Alonso, J., . . . Mokkink, L. B. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. *Qual Life Res*, 27(5), 1159-1170. doi:10.1007/s11136-018-1829-0

Thomson, W. M. (2007). Measuring change in dry-mouth symptoms over time using the Xerostomia Inventory. *Gerodontology*, 24(1), 30-35.

Thomson, W. M. (2007). Measuring change in dry-mouth symptoms over time using the Xerostomia Inventory. 24(1), 30-35. doi:10.1111/j.1741-2358.2007.00137.x

Thomson, W. M., & Williams, S. M. (2000). Further testing of the xerostomia inventory. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, 89(1), 46-50. doi:https://doi.org/10.1016/S1079-2104(00)80013-X

Trotti, A., & Eisbruch, A. (2011). Reducing xerostomia through advanced technology. *The Lancet Oncology*, 12(2), 110-111. doi:http://dx.doi.org/10.1016/S1470-2045(11)70009-2