HHS Public Access

Patient-Reported Outcome Measures in Lymphedema: A Systematic Review and COSMIN Analysis

Louise Marie Beelen, MD[1,2], **Anne-Margreet van Dishoeck, PhD**[2], **Elena Tsangaris, PhD**[1], **Michelle Coriddi, MD**[3], **Joseph H. Dayan, MD**[3], **Andrea L. Pusic, MD, MHS**[1], **Anne Klassen, PhD**[4], **Dalibor Vasilic, MD, PhD**[2]

[1]Department of Surgery, Patient-Reported Outcome, Value and Experience (PROVE) Center, Harvard Medical School, Brigham and Women's Hospital, Boston, MA;

[2]Department of Plastic and Reconstructive Surgery, Erasmus MC University Medical Center, Rotterdam, The Netherlands;

[3]Memorial Sloan Kettering Cancer Center, New York, NY;

[4]McMaster University, Hamilton, ON, Canada

Abstract

Background.—Lymphedema is a chronic and debilitating condition that affects many cancer survivors. Patient-reported outcome measures (PROMs) can give valuable insight into the impact of lymphedema on a patient's quality of life and can play an essential role in treatment decisions. This study aims to (1) identify PROMs used to assess health-related quality of life (HRQoL) in patients with lymphedema; and (2) assess the quality of the lymphedema-specific PROMs.

Methods.—We performed a systematic search to identify articles on lymphedema, quality of life, and PROMs. An overview was created of all PROMs used to assess HRQoL in patients with lymphedema. The methodological quality of the lymphedema-specific PROMs was assessed using the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) criteria.

Results.—A total of 235 articles met the inclusion criteria, of which 200 described studies using one or more PROMs as an outcome measure in patients with lymphedema. The other 35 studies described the development and/or validation of a lymphedema-specific PROM. The COSMIN assessment demonstrated that none of these PROMs met all quality standards for development.

Conclusion.—The use of PROMs in lymphedema is increasing; however, based on our findings, we cannot fully support the use of any of the existing instruments. A well-developed lymphedema-specific PROM, based on patient input, is needed to gain better insight into the impact of this condition, and can be used to measure the effect of possible medical and surgical treatments.

Lymphedema is an increasing health problem that affects up to 250 million people worldwide.[1] This condition is manifested by visible swelling, progressively decreasing function, pain, and recurrent skin infections. In Western countries, the majority of patients with lymphedema suffer from secondary lymphedema caused by cancer treatment, such as lymphadenectomy or radiotherapy.[2] Secondary lymphedema most frequently affects the upper extremities following breast cancer, the lower extremities following gynecological or urinary tract cancer, and the head and neck region following cancer treatment in this area.[2] As treatment options for cancer improve, survival rates and life expectancy continue to increase,[3] therefore more people are living longer with the adverse effects of cancer treatment, including lymphedema.

Lymphedema is a chronic and progressive condition that affects both physical and psychological health and social well-being, and may result in decreased health-related quality of life (HRQoL). Previous research in patients with breast cancer showed that patients with lymphedema report a significantly lower quality of life compared with those without lymphedema.[4,5] Patients with lymphedema may experience limited function, pain, anxiety of future progression, infections, and avoidance of activities they enjoy.

For this reason, improving HRQoL is an integral goal of lymphedema treatment. Traditionally, lymphedema is treated with conservative therapy aimed only at slowing progression. Recently, new treatment options for lymphedema have been developed, such as lymphovenous bypass and vascularized lymph node transplantation,[6,7] that may potentially reverse the condition and improve quality of life for the patient. However, no consensus has been reached on the effectiveness of these new treatment options compared with the traditional approach because the methods currently used to measure outcomes, such as circumferential measurement or limb volume, are not sufficiently reliable.[8,9]

Given the impact that lymphedema has on how a patient functions and feels, patient-reported outcome measures (PROMs) may prove to be of critical value in assessing lymphedema care and treatment. PROMs are questionnaires that provide valuable and comprehensive insights into the impact a condition has on a patient from the patient's perspective.[10] A well-developed, valid, and reliable disease-specific PROM can be used to measure important concepts of a specific condition, as well as clinical change over time. PROMs are also valuable in evaluating the effectiveness of treatment.

In recent years, there has been a sharp increase in the number of newly developed PROMs. Previously conducted reviews have focused on the outcomes and utilization of PROMs; however, there has not been a rigorous assessment of the quality of the development and content for these PROMs.[11] The purpose of this systematic literature review was to (1) provide an overview of the PROMs used to measure HRQoL in patients with lymphedema; and (2) objectively assess the development and content of the lymphedema-specific PROMs, using the Consensus-based Standards for the Selection of Health Measurement instruments (COSMIN) methodology.

METHODS

This systematic literature review was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement.[12]

Search Strategy

A systematic search of the literature was performed to find articles relating to lymphedema, quality of life, and patient-reported outcomes. The search strategy was designed with the assistance of our institution's reference librarian. The search was conducted in EMBASE, MEDLINE, Web of Science, Cochrane Central, CINAHL (EBSCOhost), PsycINFO, Ovid, and Google Scholar, from inception of each database until September 2019. Results were limited to articles written in the English language. Letters, editorials, and conference abstracts were excluded. The full search strategies can be found in the electronic supplementary material. The reference lists of relevant articles were examined to find additional articles.

Study Selection

Two reviewers (AvD and LB) independently examined the articles and screened the titles and abstracts for eligibility using predetermined inclusion and exclusion criteria. The following inclusion criteria were applied: (1) the study cohort included patients with lymphedema; (2) the study used a multidimensional PROM measuring aspects of HRQoL; and/or (3) the study described the development and/or validation of a lymphedema-specific PROM. Articles were excluded when an ad hoc instrument was used (without a proper development or validation process) or when only one health domain was measured (e.g. pain, function).

Data Extraction

The full-text versions of the potentially included articles were reviewed using a data extraction sheet with the following predetermined variables: PROM(s) used, type of patients, sample size, and whether or not the article was aimed at the development and/or validation of a PROM. The articles describing the development and/or validation of a PROM were selected for the COSMIN quality assessment.

Methodological Quality Assessment

Studies aimed at the development and/or validation of a lymphedema-specific PROM were selected for further assessment following the COSMIN criteria.[13,14] These studies had to comprise original data on one or more measurement properties of the PROM as defined in the COSMIN taxonomy.

The COSMIN criteria were developed by Terwee et al.[14] as a framework to evaluate the methodological quality of PROM development and nine measurement properties, including content validity, structural validity, internal consistency, cross-cultural validity, reliability, measurement error, criterion validity, hypothesis testing, and responsiveness. Each measurement property is rated based on standards of design requirements and preferred statistical methods. A 4-point scoring system is used to rate each standard as 'very

good', 'adequate', 'doubtful' or 'inadequate'. The overall rating per measurement property is determined by the lowest rating of any standard in the box.

Two independent reviewers (AvD and LB) performed the COSMIN evaluation. Discrepancies of opinion were resolved by consensus between the two reviewers or, if no consensus was obtained, with the help of a third reviewer (ET). The results of this assessment were organized according to tables provided in the COSMIN user manual.[15,16] The percentage agreement between the two reviewers performing the COSMIN evaluation was calculated by dividing the number of ratings with agreement by the total number of ratings performed in this study. Based on a study on the inter-rater agreement of the COSMIN checklist, we considered a percentage agreement above 80% appropriate.[17]

RESULTS

Search Results

A total of 4459 articles were identified through the initial search (Fig. 1), and one additional article was identified through review of citations.[18] After the removal of duplicates, 2321 articles were screened on title and abstract and 399 articles were selected for full-text assessment. Among these articles, 235 were included in the study. Most articles ($n = 200$) measured HRQoL using at least one PROM in a lymphedema study population. The remaining 35 articles described the development and/or validation of a lymphedema-specific PROM.

Overview of Health-Related Quality-of-Life Patient-Reported Outcome Measures (PROMs) Being Used in Patients with Lymphedema

Fifty-four different PROMs were used in the 200 articles assessing HRQoL in patients with lymphedema, most of which (55%) used two or more PROMs. The PROMs used were divided into the following groups: generic ($n = 26$), oncology-specific ($n = 16$), and lymphedema-specific ($n = 12$).

An overview of the identified PROMs and the number of times they were used can be found in electronic supplementary Table 1. The most frequently used generic PROM was the SF-36 ($n = 54$ studies) and the most frequently used oncology-specific PROMs were the EORTC QLQ-C30 ($n = 44$ studies) and the EORTC QLQ-BR12 ($n = 26$ studies). Forty-six studies used one or more lymphedema-specific PROMs, with the LYMQOL being used most frequently ($n = 26$ studies).

Development and Validation of Lymphedema-Specific PROMs

Our literature search yielded articles describing the development and/or validation of 17 lymphedema-specific PROMs. The two lymphedema-specific PROMs for which we did not find a development or validation article through our search were the Instituto Dermopatico Dell Immacolata – Italian Lymphedema Association (IDI-ILA) and the Wesley Clinic Lymphedema Scale (WCLS). The IDI-ILA was developed specifically for patients with melanoma and was only available in Italian.[19] The WCLS was developed by adapting the Functional Living Index–Cancer questionnaire, by replacing the words 'illness' or 'cancer'

with 'lymphedema',[20] and was therefore considered an ad hoc instrument. It was decided to not include these two PROMs in the quality assessment.

For 7 of the 17 PROMs, we found articles describing their development and/or validation, but we identified no studies using the PROM in a clinical study.[21–27] We did however decide to include these PROMs in our COSMIN quality analysis because they met our study criteria. This explains why the list of lymphedema-specific PROMs used in clinical studies (electronic supplementary Table 1c) differs from the list of PROMs that underwent COSMIN analysis (Table 1).

Overall, 35 articles describing the development or validation of 17 lymphedema-specific PROMs were eligible for assessment using the COSMIN checklist. The articles included 13 original validation studies,[18,22,24,26,28–36] four adaptations or revisions of existing PROMs,[23,27,37,38] and 18 translations of PROMs into other languages.[21,25,39–54] Table 1 provides an overview of the assessed PROMs and their characteristics. The percentage agreement between the two reviewers performing the COSMIN evaluation was 91%.

1. *PROM Development*: The PROM development process was published for 13 of the 17 lymphedema-specific PROMs. For the four other PROMS, the development process was either presented at a conference but not published (BCLE-SEI,[55] LYMQOL[56]) or the development process was not published ([S]LQOLI,[25] LyQLI[38]).

An overview of the quality of PROM development ratings can be found in Table 2. Based on the COSMIN criteria and 4-point rating scale, we found that the quality of the development process was rated as doubtful for three PROMs and rated as inadequate for nine PROMS.

Part of the development process is concept elicitation, defined as "the process by which concepts (e.g. symptoms and impacts) that are important to patients emerge spontaneously through the use of open-ended questions in an interview setting".[57] This part was rated inadequate or doubtful for all PROMS. Only one of the investigated PROMs, the ULL-27, used patient interviews for concept elicitation. An overview of the methods used for item generation and reduction per PROM can be found in Table 4.

Pilot tests or cognitive interview studies had been performed for seven of the instruments. The criteria for the cognitive interview study included testing the comprehensibility, conceptualizing clarity, and the comprehensiveness, as a measure of completeness. Preferred methods to assess comprehensibility are, for example, the think-aloud method, or other forms of cognitive interviewing.[14,58] Comprehensibility was assessed in five studies, but all used methods of inadequate or doubtful quality and gave limited insight into the retrieved data and possible adaptations. Comprehensiveness was tested in six studies and showed doubtful methodological quality. Reasons to grade the quality as doubtful were, for example, questions not tested in the final form, no appropriate method (e.g. cognitive interview) used, or unclear data analysis.

2. *Content Validity*: Content validity is defined as "the degree to which the content of an instrument measures the construct(s) it purports to measure",[59] and is considered to be the most important measurement property of a PROM. In the

COSMIN guidelines, Terwee et al.[12] describe three aspects of content validity: (1) relevance (the items of the PROM are relevant for the construct of interest within the specific population and context of use); (2) comprehensiveness (all key items are included); and (3) comprehensibility (patients understand all items as intended). It is recommended that these three aspects are assessed in a study by asking both patients and professionals. Of the 34 studies we assessed, 10 described content validity. None of these articles used adequate methods to assess content validity and were therefore rated poorly (Table 3).

3. *Measurement Properties*: Subsequently, we rated the measurement properties of the PROMs in terms of reliability, validity, and responsiveness. An overview of the summary scores for the methodological quality of each measurement property can be found in Table 3. The LSIDS-H&N could not be evaluated on these terms, since only the development process of this PROM was published and no information about its measurement properties was found.[30] None of the studies reported on cross-cultural validity, and very few studies ($n = 8$) reported on the responsiveness of the PROM. The studies that did test responsiveness used methods of 'doubtful' or 'inadequate' quality, with one exception (LyQLI).[36]

Structural validity is defined as "the degree to which the scores of a PROM are an adequate reflection of the dimensionality of the construct to be measured".[12] This measurement property is rated on statistical method (e.g. factor analysis is recommended for studies that take a Classical Test Theory approach) and sample size. Structural validity was rated for 18 studies, with scores varying from 'inadequate' ($n = 7$ studies) to 'very good' ($n = 1$ study).

Internal consistency, defined as "the degree of the interrelatedness among the items",[12] is usually assessed using Cronbach's alpha. This was determined with high methodological quality in almost all studies, with 32 studies being rated 'very good' (Table 3).

Reliability is defined as "the degree to which the measure is free from measurement error"[12] and can be tested with repeated measurements in stable patients, with an appropriate time interval, and under similar test conditions. The methods used to determine reliability in the articles we assessed were scored as 'inadequate' for 24 studies, mostly because the repeated measurements were conducted in a suboptimal manner (e.g. patients were not stable, inappropriate time interval, different test conditions) or because inadequate statistical methods had been used.

Measurement error, "the systematic and random error of a patient's score that is not attributed to changes in the construct to be measured",[12] was rated 'inadequate' in all of the nine articles describing this measurement property. This rating was often based on suboptimal repeated measurement conditions and inadequate use of statistical methods.

Criterion validity is defined as "the degree to which the scores of a PROM are an adequate reflection of a gold standard". It was rated in 11 studies and the quality ranged from 'inadequate' ($n = 3$) to 'very good' ($n = 2$). Construct validity (or hypotheses testing) was most often measured by comparison of the PROM with other outcome measures (convergent

validity, $n = 28$ studies) and was carried out with 'adequate' methodological quality in 16 studies.

    4.   *Criteria for Good Measurement Properties*: The results of each study on the different measurement properties were extracted and are shown in electronic supplementary Table 5. Each result was rated as sufficient (?), insufficient (−) or indeterminate (?) following COSMIN criteria for good measurement properties. While the previous ratings related to the methodological quality of studies on measurement property, the criteria for good measurement properties refer to the quality of the PROM itself. The results on the different measurement properties show multiple results of indeterminate quality. Furthermore, we see multiple PROMS scoring insufficiently on criterion validity and responsiveness. Because of the limited number of validation studies per PROM, we chose not to summarize the results and thus to not grade the total level of evidence per PROM.

Following the COSMIN methodology, we were unable to formulate a recommendation on the most suitable PROM, as the number of validation studies was limited and the methodological quality of multiple studies did not meet the COSMIN standards.

DISCUSSION

In this systematic review, we have provided an overview of the various PROMs used to measure HRQoL in patients with lymphedema. We found a large number of different PROMs used to measure HRQoL in this population. This breadth of PROMs implies that there is a lack of consensus on the most suitable PROM. Moreover, this variety or heterogeneity of PROMs makes it practically impossible to compare outcomes between studies. This poses limitations for international research efforts aimed at improving treatment methods and HRQoL in patients with lymphedema. Additionally, we found that several studies used a generic PROM, although previous studies have shown that generic PROMs do not adequately capture disease-specific concerns and are not suitable to measure the effect of treatment.[60–63]

The second aim of this review was to assess the development and psychometric properties of existing lymphedema-specific PROMs. To this end, we applied the COSMIN methodology. In our assessment of the current lymphedema-specific instruments, we discovered that no published information exists on the development process of four published lymphedema PROMs. Moreover, none of the lymphedema-specific PROMs met all the COSMIN quality standards for development. A major shortcoming in the development process was the lack of patient involvement, which is an essential aspect of the development of a PROM.[64,65] Patient involvement is crucial to the development of PROMs that measure outcomes that matter to patients. Widely recommended methods for concept elicitation include individual interviews and focus groups.[65–67] Half of the lymphedema-specific PROMs we examined did not include any patient input in the development phase. Of the seven PROM development studies that did include a form of patient input, only one study conducted qualitative interviews with patients (ULL-27).[18] The other six studies used a patient survey, which may fail to adequately capture the patient's perceptions, feelings, and viewpoints.

Rigorous development of a PROM is a challenging and time-intensive process; however, it is a vital step to create an adequate instrument.[65] Unfortunately, we found that most studies have paid insufficient attention to the development phase or provided insufficient information on their steps taken. A disease-specific PROM, able to capture the outcome as experienced by the patient, would be highly valuable for patients with lymphedema and caregivers.

The overall methodology of the lymphedema-specific PROM validation studies was found to be of low to moderate quality (as demonstrated in Table 3). The PROMS that showed the best methodological quality include the LISS, Lymph-ICF, Lymph-ICF-LL, PBI-L, and ULL-27. Most validation studies did not report on the responsiveness of the instrument, even though responsiveness is considered one of the major advantages of a disease-specific PROM over a generic PROM. If an instrument has a poor ability to capture change (responsiveness), it can result in false-negative outcomes on the effect of treatment.[68] However it must be noted that a number of PROMs have not been used in research beyond the development phase, and it is possible that future studies reporting on responsiveness are forthcoming.

In the literature, we identified two review articles describing an assessment of available lymphedema PROMs. Pusic et al. reviewed PROMs evaluating QoL in patients with breast cancer-related lymphedema (BCRL). They identified two lymphedema-specific PROMs and found the ULL-27 to have the strongest psychometric properties.[69] Their findings are in line with our current assessment. Cornelissen et al. also performed a review of QoL PROMs in patients with BCRL and found the Lymph-ICF and LyQLI to be the most complete and accurate PROMs;[70] however, that study did not assess the development process of the PROMs and did not include measurement properties other than reliability (Cronbach alpha coefficient) in their review. The difference in findings in these two reviews emphasizes the importance of using standardized criteria to rate PROMs, such as the COSMIN checklist, which was developed in a Delphi study involving 158 experts from 21 countries.[14] A recent study by Coriddi et al. provides an overview of PROMs used in the evaluation of the surgical treatment of lymphedema.[11] These authors found that a variety of PROMs were used, demonstrating a lack of consensus among lymphedema researchers and the need for a critical appraisal of the development and validation of the lymphedema-specific PROMs.

The limited ability to measure an outcome following any medical or surgical intervention for lymphedema makes it exceedingly difficult to advance this field towards a cure. Limb volume and patient-reported outcomes are the two most frequently used metrics, neither of which are highly reliable. Limb volume is dynamic, can be manipulated by decongestive therapy following any surgical or medical intervention, and is meaningless in a patient with minimal volume difference. Consequently, there has been more recent focus on patient-reported outcomes. However, the results in this study demonstrate that the currently available instruments are inadequate and may not represent the totality of the patient's quality of life. The lack of significant open-ended patient input during the development phase introduces bias into what is being measured. For example, most, if not all, of these questionnaires were developed by or in collaboration with lymphedema therapists, with little patient input.

CONCLUSION

Lymphedema is an important and growing health problem that negatively effects HRQoL. Currently, a number of lymphedema-specific PROMs are available for use, but most without evidence of an adequate development process and none that met methodological quality standards. Therefore, based on our findings in this review, we recommend the development of a new PROM based on extensive qualitative input from patients with lymphedema, and adequately validated in studies showing good methodological quality.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENT

REFERENCES

1. Földi M, Földi E, Strößenreuther R, Kubik S. Földi's textbook of lymphology: for physicians and lymphedema therapists. Elsevier Health Sciences; 2012.

2. Rockson SG, Rivera KK. Estimating the population burden of lymphedema. Ann N Y Acad Sci. 2008;1131(1):147–154. [PubMed: 18519968]

3. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA: A Cancer J Clin. 2019;69(1):7–34.

4. Penha TRL, Botter B, Heuts EM, Voogd AC, von Meyenfeldt MF, van der Hulst RR. Quality of life in patients with breast cancer-related lymphedema and reconstructive breast surgery. J Reconstr Microsurg. 2016;32(06):484–90. [PubMed: 26919383]

5. Pusic AL, Cemal Y, Albornoz C, et al. Quality of life among breast cancer patients with lymphedema: a systematic review of patient-reported outcome instruments and outcomes. J Cancer Surviv. 2013;7(1):83–92. [PubMed: 23212603]

6. Chang DW, Masia J, Garza Iii R, Skoracki R, Neligan PC. Lymphedema: surgical and medical therapy. Plast Reconstr Surg. 2016;138(3S):209S–18S. [PubMed: 27556764]

7. Schaverien MV, Coroneos CJ. Surgical Treatment of Lymphedema. Plast Reconstr Surg. 2019;144(3):738–58. [PubMed: 31461041]

8. Hayes SC, Janda M, Cornish BH, Battistutta D, Newman B. Lymphedema secondary to breast cancer: how choice of measure influences diagnosis, prevalence, and identifiable risk factors. Lymphology. 2008;41(1):18–28. [PubMed: 18581955]

9. Cornish BH, Chapman M, Thomas BJ, Ward LC, Bunce IH, Hirst C. Early diagnosis of lymphedema in postsurgery breast cancer patients. Ann N Y Acad Sci. 2000;904(1):571–5. [PubMed: 10865807]

10. Black N Patient reported outcome measures could help transform healthcare. BMJ. 2013;346:f167. [PubMed: 23358487]

11. Coriddi M, Dayan J, Sobti N, et al. Systematic review of patient-reported outcomes following surgical treatment of lymphedema. Cancers. 2020;12(3):565.

12. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med. 2009;151(4):264–9, W264. [PubMed: 19622511]

13. Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147–57. [PubMed: 29435801]

14. Terwee CB, Prinsen CAC, Chiarotto A, et al. COSMIN methodology for evaluating the content validity of patient-reported outcome measures: a Delphi study. Qual Life Res. 2018;27(5):1159–70. [PubMed: 29550964]

15. Mokkink LB, De Vet HCW, Prinsen CAC, et al. COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1171–9. [PubMed: 29260445]

16. Mokkink LB, Prinsen C, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs): user manual. Version 1.0 2018. p. 78.

17. Mokkink LB, Terwee CB, Gibbons E, 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):82. [PubMed: 20860789]

18. Launois R, Megnigbeto AC, Pocquet K, Alliot F. A specific quality of life scale in upper limb lymphedema: the ULL-27 questionnaire. Lymphology. 2002;35(1–760):181–7.

19. Carmeli E, Bartoletti R. Retrospective trial of complete decongestive physical therapy for lower extremity secondary lymphedema in melanoma patients. Supportive Care Cancer. 2011;19(1):141–7.

20. Mirolo BR, Bunce IH, Chapman M, et al. Psychosocial benefits of postmastectomy lymphedema therapy. Cancer Nurs. 1995;18(3):197–205. [PubMed: 7600551]

21. Shi S, Lu Q, Fu MR, et al. Psychometric properties of the Breast Cancer and Lymphedema Symptom Experience Index: The Chinese version. Eur J Oncol Nurs. 2016;20:10–6. [PubMed: 26071198]

22. Augustin M, Bross F, Földi E, Vanscheidt W, Zschocke I. Development, validation and clinical use of the FLQA-l, a disease-specific quality of life questionnaire for patients with lymphedema. Vasa J Vasc Dis. 2005;34(1):31–5.

23. Weiss J, Daniel T. Validation of the lymphedema life impact scale version 2: a condition-specific measurement tool for persons with lymphedema. Rehabil Oncol. 2018;36(1):28–36.

24. Blome C, Augustin M, Heyer K, et al. Evaluation of patient-relevant outcomes of lymphedema and lipedema treatment: Development and validation of a new benefit tool. Eur J Vasc Endovasc Surg. 2014;47(1):100–7. [PubMed: 24239143]

25. Klernäs P, Kristjanson LJ, Johansson K. Assessment of quality of life in lymphedema patients: Validity and reliability of the Swedish version of the Lymphedema Quality Of Life Inventory (LQOLI). Lymphology. 2010;43(3):135–45. [PubMed: 21226416]

26. Williams AE, Rapport F, Russell IT, Hutchings HA. Psychometric development of the Upper Limb Lymphedema Quality of Life Questionnaire demonstrated the patient-reported outcome measure to be a robust measure for breast cancer–related lymphedema. J Clin Epidemiol. 2018;100:61–70. [PubMed: 29723583]

27. De Vrieze T, Vos L, Gebruers N, et al. Revision of the Lymphedema Functioning, Disability and Health Questionnaire for Upper Limb Lymphedema (Lymph-ICF-UL): reliability and validity. Lymphatic Res Biol. 2019;17(3):347–55.

28. Weiss J, Daniel T. Validation of the lymphedema life impact scale (LLIS): a condition-specific measurement tool for persons with lymphedema. Lymphology. 2015;48(3):128–38. [PubMed: 26939160]

29. Ridner SH, Dietrich MS. Development and validation of the Lymphedema Symptom and Intensity Survey-Arm. Supportive Care Cancer. 2015;23(10):3103–3112.

30. Deng J, Ridner SH, Murphy BA, Dietrich MS. Preliminary development of a lymphedema symptom assessment scale for patients with head and neck cancer. Supportive Care Cancer. 2012;20(8):1911–8.

31. Devoogdt N, Van Kampen M, Geraerts I, Coremans T, Christiaens MR. Lymphoedema Functioning, Disability and Health questionnaire (Lymph-ICF): reliability and validity. Phys Ther. 2011;91(6):944–57. [PubMed: 21493748]

32. Devoogdt N, De Groef A, Hendrickx A, et al. Lymphoedema Functioning, Disability and Health Questionnaire for Lower Limb Lymphoedema (Lymph-ICF-LL): reliability and validity. Phys Ther. 2014;94(5):705–21. [PubMed: 24415775]

33. Keeley V, Crooks S, Locke J, Veigas D, Riches K, Hilliam R. A quality of life measure for limb lymphoedema (LYMQOL). J Lymphoedema. 2010;5(1):26–37.

34. Launois R, Alliot F. Quality of life scale in upper limb lymphedema: a validation study. Lymphology. 2000;33 Suppl:266–70.

35. Ridner SH, Doersam JK, Stolldorf DP, Dietrich MS. Development and Validation of the Lymphedema Symptom Intensity and Distress Survey-Lower Limb. Lymphatic Res Biol. 2018;16(6):538–46.

36. Klernäs P, Johnsson A, Boyages J, Brorson H, Munnoch A, Johansson K. Test of responsiveness and sensitivity of the questionnaire "lymphedema Quality of Life Inventory". Lymphatic Res Biol. 2018;16(3):300–8.

37. Augustin M, Conde Montero E, Hagenström K, Herberger K, Blome C. Validation of a short-form of the Freiburg Life Quality Assessment for lymphoedema (FLQA-LS) instrument. Br J Dermatol. 2018;179(6):1329–33. [PubMed: 29927481]

38. Klernäs P, Johnsson A, Horstmann V, Kristjanson LJ, Johansson K. Lymphedema Quality of Life Inventory (LyQLI)-Development and investigation of validity and reliability. Qual Life Res. 2015;24(2):427–39. [PubMed: 25633655]

39. De irmenci B, Tüzün , Of NS, Oral A, Sindel D. Reliability and validity of Turkish version of lymphedema life impact scale. Turk J Phys Med Rehabil. 2019;65(2):147–53. [PubMed: 31453555]

40. Haghighat S, Montazeri A, Zayeri F, Ebrahimi M, Weiss J. Psychometric evaluation of the Persian version of the Lymphedema Life Impact Scale (LLIS, version 1) in breast cancer patients. Health Qual Life Outcomes. 2018;16(1):132. [PubMed: 29954381]

41. Orhan C, Üzelpasaci E, Baran E, et al. The reliability and validity of the turkish version of the lymphedema life impact scale in patients with breast cancer-related lymphedema. Cancer Nurs. 2020;43(5):375–83. [PubMed: 30921030]

42. Abu Sharour L Psychometric evaluation of the Arabic version of the lymphedema life impact scale in breast cancer patients. Breast J. 2020;26(3):563–5. [PubMed: 31512781]

43. Kostanoglu A, Hosbay Z, Tarakci E. Lymphoedema functioning, disability and health questionnaire Turkish version: translation, cross-cultural adaptation and validation. J Phys Ther Sci. 2016;28(6):1728–32. [PubMed: 27390404]

44. Grarup KR, Devoogdt N, Strand LI. The Danish version of Lymphoedema Functioning, Disability and Health Questionnaire (Lymph-ICF) for breast cancer survivors: translation and cultural adaptation followed by validity and reliability testing. Physiother Ther Pract. 2019;35(4):327–40.

45. Kostanoglu A, Mbata GB, Gokmen GY, Uysal O. The Lymphedema Functioning, Disability, and Health Questionnaire for Lower Limb Lymphedema: Translation, reliability, and validation study of the Turkish version. Turk J Thoracic Cardiovasc Surg. 2017;25(4):586–91.

46. Wang CM, Lee SY, Hsu KF, Lin CF, Ma MC, Hsu YY. Psychometric evaluation of a Chinese version of Lymphoedema Functioning, Disability and Health Questionnaire for Lower Limb Lymphoedema in women with gynaecological cancer surgery. Eur J Cancer Care (Engl). 2018;27(6):e12940. [PubMed: 30311336]

47. Bakar Y, Tugral A, Ozdemir O, Duygu E, Uyeturk U. Translation and Validation of the Turkish Version of Lymphedema quality of life tool (LYMQOL) in patients with breast cancer related lymphedema. Eur J Breast Health. 2017;13(3):123–8. [PubMed: 28894851]

48. Borman P, Yaman A, Denizli M, Karahan S, Ö zdemir O. The reliability and validity of Lymphedema Quality of Life Questionnaire-Arm in Turkish patients with upper limb lymphedema related with Breast cancer. Turk J Phys Med Rehabil. 2018;64(3):205–12. [PubMed: 31453513]

49. Karayurt Ö, Deveci Z, Eyigör S, Özgür nbat M. Adaptation of quality of life measure for limb lymphedema-arm in Turkish Women With Breast Cancer-Related Lymphedema. Cancer Nurs. Epub 28 8 2019. 10.1097/ncc.0000000000000741

50. van de Pas CB, Biemans AAM, Boonen RSM. Validation of the Lymphoedema Quality-of-Life Questionnaire (LYMQOL) in Dutch patients diagnosed with lymphoedema of the lower limbs. Pblebology. 2016;31(4):257–63.

51. Bakar Y, Tu ral A. Translation, reliability, and validation of the Turkish version of the Lymphedema Quality-of-Life tool in Turkish-speaking patients with lower limb Lymphedema. J Vasc Nurs. 2019;37(1):11–7. [PubMed: 30954192]

52. Borman P, Yaman A, Denizli M, Karahan S. The reliability and validity of lymphedema quality of life questionnaire-leg in Turkish patients with lower limb lymphedema. Lymphat Res Biol. 2020;18(1):42–8. [PubMed: 31135275]

53. Duygu E, Bakar Y, Keser I. An important tool in lymphedema management: validation of Turkish version of the patient benefit index-lymphedema. Lymphat Res Biol. 2020;18(1):49–55. [PubMed: 31145018]

54. Viehoff PB, Van Genderen FR, Wittink H. Upper limb lymphedema 27 (ULL27): Dutch translation and validation of an illness-specific health-related quality of life questionnaire for patients with upper limb lymphedema. Lymphology. 2008;41(3):131–8. [PubMed: 19013881]

55. Fu MR, Cleland CM, Kang Y. Measuring lymphedema symptom burdens: a psychometric study. Paper presented at the The Multinational Association of Supportive Care in Cancer's Annual Meeting (MASCC/ISOO): New York, 2012.

56. Keeley VL, Veigas D, Crooks S, Locke J, Forrow H. The development of a condition-specific quality of life measure for lymphoedema (LYMQOL). Eur J Lymphol. 2004;12(41):36.

57. Consortium YHE. Concept Elicitation. 2016. Available at: https://yhec.co.uk/glossary/concept-elicitation/.

58. Jobe JB, Mingay DJ. Cognitive research improves questionnaires. Am J Public Health. 1989;79(8):1053–5. [PubMed: 2751028]

59. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. J Clin Epidemiol. 2010;63(7):737–45. [PubMed: 20494804]

60. Pusic AL, Chen CM, Cano S, et al. Measuring quality of life in cosmetic and reconstructive breast surgery: a systematic review of patient-reported outcomes instruments. Plast Reconstr Surg. 2007;120(4):823–37. [PubMed: 17805107]

61. Guyatt GH, Bombardier C, Tugwell PX. Measuring disease-specific quality of life in clinical trials. CMAJ. 1986;134(8):889. [PubMed: 3955482]

62. Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. Med Care. 1989;27(3 Suppl):S217–32. [PubMed: 2646490]

63. Wiebe S, Guyatt G, Weaver B, Matijevic S, Sidwell C. Comparative responsiveness of generic and specific quality-of-life instruments. J Clin Epidemiol. 2003;56(1):52–60. [PubMed: 12589870]

64. US FDA. Patient reported outcome measures: Use in medical product development to support labeling claims. Guidance for Industry. US FDA; 12 2009.

65. Lasch KE, Marquis P, Vigneux M, et al. PRO development: rigorous qualitative research as the crucial foundation. Qual Life Res. 2010;19(8):1087–96. [PubMed: 20512662]

66. Patrick DL, Burke LB, Gwaltney CJ, et al. Content validity—establishing and reporting the evidence in newly developed patient-reported outcomes (PRO) instruments for medical product evaluation: ISPOR PRO good research practices task force report: part 1—eliciting concepts for a new PRO instrument. Value Health. 2011;14(8):967–77. [PubMed: 22152165]

67. Turner RR, Quittner AL, Parasuraman BM, Kallich JD, Cleeland CS, Mayo FDAP-ROCMG. Patient-reported outcomes: instrument development and selection issues. Value Health. 2007;10 Suppl 2:S86–93. [PubMed: 17995478]

68. Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (eds). Cochrane Handbook for Systematic Reviews of Interventions version 6.0 2019; www.training.cochrane.org/handbook. Accessed 2 Oct 2019.

69. Pusic AL, Cemal Y, Albornoz C, et al. Quality of life among breast cancer patients with lymphedema: a systematic review of patient-reported outcome instruments and outcomes. J Cancer Survivorship. 2013;7(1):83–92.

70. Cornelissen AJM, Kool M, Keuter XHA, et al. Quality of life questionnaires in breast cancer-related lymphedema patients: review of the literature. Lymphat Res Biol. 2018;16(2):134–9. [PubMed: 29437542]
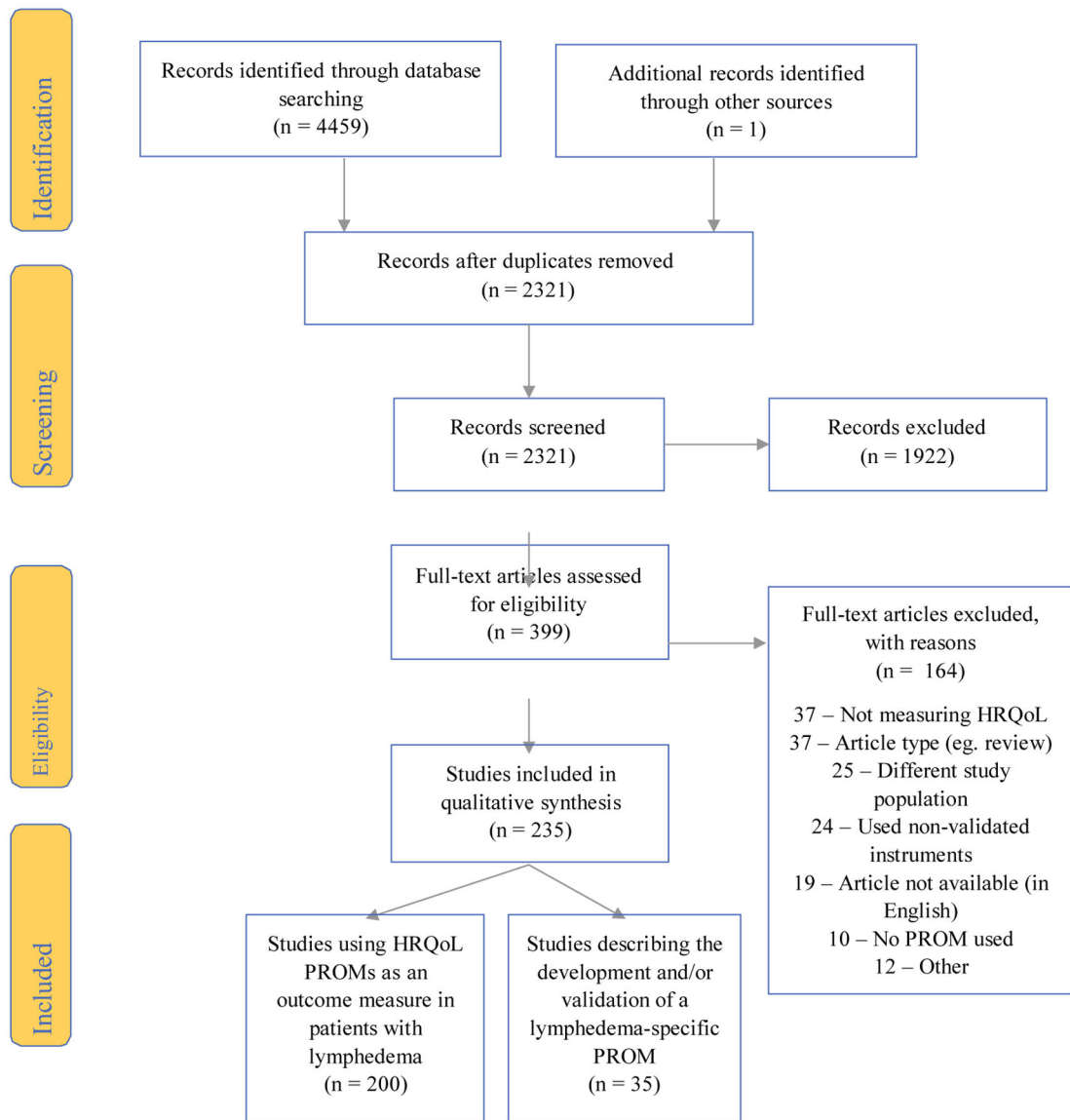
FIG. 1.
PRISMA flow diagram for systematic review. *HRQoL* health-related quality of life, *PROM* patient-reported outcome measure

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

TABLE 1

Characteristics of lymphedema-specific PROMs

PROM (first article)	Target population	No. of items	Subscales	Response options	Scoring	Original language	Available translations
BCLE-SEI (Fu et al.,2012)[55]	Patients with breast cancer-related lymphedema	34	Symptoms, Distress	Yes/no questions, 5-point scale	Symptom score, distress score	English	Chinese
FLQA-I (Augustin et al., 2005)[22]	Patients with all forms of lymphedema	92	Physical Complaints, Everyday Life, Social Life, Emotional Status, Treatment, Satisfaction, Profession/ Household	Unknown	Unknown	German	
FLQA-LS (Augustin et al., 2018)[37]	Patients with all forms of lymphedema	33	Physical Impairments, Daily Life, Social Life, Psychological Well-Being, Therapy of the Lymph Condition, Global Assessment	5-point scale, 11-point visual analog scale	Unknown	German	English
LLIS (Weiss and Daniel, 2015)[28]	Patients with any extremity lymphedema	18	Physical, Psychosocial, Functional	5-point scale	Total score, subscale score	English	
LLIS 2.0 (Weiss and Daniel, 2018)[23]	Patients with any extremity lymphedema	18	Physical, Psychosocial, Functional, Infection Occurrence	5-point scale	Total score, subscale score	English	
LSIDS-A (Ridner and Dietrich, 2015)[29]	Patients with upper extremity lymphedema	30	Soft Tissue Sensation, Neurologic Sensation, Function, Biobehavioral, Resource, Sexuality, Activity	Yes/no questions, 10-point scale	Total score, subscale score, intensity and distress scores	English	
LSIDS-H&N (Deng et al., 2012)[30]	Patients with head and neck cancer-related lymphedema	67	Altered Sensations, General Psychosocial Issues, Functional Impairments and Symptom Burden, Anatomical Sites of Swelling	Yes/no questions, 10-point scale, 11-point scale	Unknown	English	
LSIDS-L (Ridner et al., 2018)[35]	Patients with lower extremity lymphedema	31	Activity, Soft Tissue Sensation, Pain, Resources, Biobehavioral, Neurologic Sensation, Function, Sexuality	Yes/no questions, 5-point scale	Total score, subscale score	English	
Lymph-ICF (Devoogdt et al., 2011)[31]	Patients with breast cancer-related lymphedema	29	Physical Function, Mental Function, Household Activities, Mobility Activities, Life and Social Activities	Visual analog scale	Total score, subscale score	Dutch	English, Turkish, Danish
Lymph-ICF-UL (De Vrieze et al., 2019)[27]	Patients with upper extremity lymphedema	29	Physical Function, Mental Function, Household Activities, Mobility Activities, Life and Social Activities	11-point scale	Total score, subscale score	Dutch	English, French
Lymph-ICF-LL (Devoogdt et al., 2014)[32]	Patients with lower extremity lymphedema	28	Physical Function, Mental Function, General Tasks/Household Activities, Mobility Activities, Life Domains/ Social Life	11-point scale	Total score, subscale score	Dutch	English, Turkish, Portuguese
LYMQOL (Keeley et al., 2010)[33]	Patients with any extremity lymphedema	Arm: 28 Leg: 27	Function, Appearance, Symptoms, Mood, Overall QoL	4-point scale, 11-point scale, open question	Total score, subscale score	English	Dutch, Turkish
LyQLI (Klernäs et al., 2015)[38]	Patients with all forms of lymphedema	45	Physical, Psychosocial, Practical, General HRQoL	4-point scale, yes/no question	Subscale score	Swedish	English

PROM (first article)	Target population	No. of items	Subscales	Response options	Scoring	Original language	Available translations
S)LQOLI (Klernas et al., 2010)[25]	Patients with all forms of lymphedema	58	Physical, Emotional, Social, Practical, General QoL	point scale, 11-point scale	Unknown	English (LQOLI), Swedish (SLQOLI)	
PBI-L (Blome et al., 2014)[24]	Patients with lymphedema and/or lipedema	46	Patient Needs, Patient Benefit	5-point scale	Total score	German	English, Turkish
ULL-27 (Launois and Alliot, 2000)[34]	Patients with breast cancer-related lymphedema	27	Physical, Psychological, Social	5-point scale	Total score, subscale score	French	Dutch
ULLQoL (Williams et al., 2018)[26]	Patients with breast cancer-related lymphedema	14	Physical Well-Being, Emotional Well-Being	5-point scale	Total score, subscale score	English	

PROM patient-reported outcome measure, *LQOLI* Lymphedema Quality of Life Inventory, *SLQOLI* Swedish version of the Lymphedema Quality of Life Inventory, *QoL* quality of life

TABLE 2

Methodological quality of PROM development

PROM	PROM design — General design requirements					Concept elicitation	Total PROM design	CI study — CI study performed in the sample representing the target population	Comprehensibility	Comprehensiveness	Total CI study	Total PROM development
	Clear construct	Clear origin of construct	Clear target population for which the PROM was developed	Clear context of use	PROM developed in the sample representing the target population							
BCLE-SEI												
FLQA-I	I	D	I	D	A	D	I	I				I
FLQA-LS[a]	I	D	I	D	I	I	I					I
LISS	I	I	I	D	I	D	I	D		D	D	I
LISS 2.0[a]	I	I	I	D	I	D	I	I				I
LSIDS-A	V	V	I	V	A	I	I	V	D	D	D	I
LSIDS-H&N[a]	V	D	V	V	V	I	I	V	D	D	D	I
LSIDS-L[a]	V	D	V	V	V	I	I	I				I
Lymph-ICF	V	V	V	V	V	D	D	V	D	D	D	D
Lymph-ICF-UL[a]								V	D	A	D	D
Lymph-ICF-LL	V	V	V	V	V	D	D	V	D	D	D	D
LYMQOL												
LyQLI[a]												
BPI-L	V	D	V	V	V	I	I	V	I	D	I	I
(S)LQOLI[a]												
ULL-27	V	D	V	V	A	D	D	I				I
ULLQoL	V	D	V	D	V	D	D	V				D

Empty cells indicate that no information was available on this item

PROM patient-reported outcome measure, *CI* cognitive interview, *V* very good, *A* adequate, *D* doubtful, *I* inadequate

[a]This PROM is an adaptation or revision of a previously developed PROM

TABLE 3

Quality of studies on measurement properties

PROM	Article	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity		Responsiveness			
								Convergent validity	Known groups validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups	Comparison before and after intervention
BCLE-SEI	Shi et al., 2016[21]	A	V		I		D	D	D				
FLQA-I	Augustin et al., 2005[22]		V		I			A	I				I
FLQA-LS	Augustin et al., 2018[37]	I	V					D					
LISS	Weiss and Daniel, 2015[28]		V		I		V	D	D				D
	Degirmenci et al., 2019[39]	D	V		I			D					
	Haghighat et al., 2018[40]	D	V		I			A	I				
LLIS 2.0	Weiss and Daniel, 2018[23]		V		I	I	I	A					
	Orhan et al., 2020[41]	D	D		I		D		I				
	Abu Sharour, 2020[42]	I	V					D					
LSIDS-A	Ridner and Dietrich, 2015[29]		V		I			A					
LSIDS-H&N	Deng et al., 2012[30]												
LSIDS-L	Ridner et al., 2018[35]		V					D	D				

PROM	Article	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity		Responsiveness			
								Convergent validity	Known groups validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups	Comparison before and after intervention
Lymph-ICF	Devoogdt et al., 2011[31]		V		I	I		A	A				
	Kostanoglu et al., 2016[43]		V		I	I	I	A					
	Grarup et al., 2019[44]		V		I	I							
Lymph-ICF-UL	De Vrieze et al., 2019[27]		V		I	I		A					
Lymph-ICF-LL	Devoogdt et al., 2014[32]		V		I	I		A					
	Kostanoglu et al., 2017[45]		V		I			A					
LYMQOL	Wang et al., 2018[46]	I	V		I		D						
	Keeley et al., 2010[33]	I	V		I	I	D	I					D
LYMQOL-arm	Bakar et al., 2017[47]	I	V		D	I	I	D					
	Borman et al., 2018[48]	A	V		I			V					
	Karayurt et al., 2019[49]	A	V						I				
LYMQOL-leg	van de Pas et al., 2016[50]	I	V		I	I		A					
	Bakar and Tural, 2019[51]	I	V		D		V						
	Borman et al., 2020[52]	A	V		I			A	D				

PROM	Article	Structural validity	Internal consistency	Cross-cultural validity	Reliability	Measurement error	Criterion validity	Construct validity		Responsiveness			
								Convergent validity	Known groups validity	Comparison with gold standard	Comparison with other instruments	Comparison between subgroups	Comparison before and after intervention
LyQLI	Klemäs et al., 2015[38]	D	V		I			A					
	Klemäs et al., 2018[36]										A		V
PBI-L	Blome et al., 2014[24]	A	V					A					I
	Duygu et al., 2020[53]		V		I			D					
(S)LQOLI	Klemäs et al., 2010[25]				D			A					
ULL-27	Launois and Alliot 2000[34]	I	V		I			D			I	I	
	Launois et al., 2002[18]	V	V		I		V	A	D			D	
	Viehoff et al., 2008[54]		V				D	D	D				
ULLQoL	Williams et al., 2018[26]	A	V		I			A				D	

Empty cells indicate that no information was available on this item

PROM patient-reported outcome measure, *V* very good, *A* adequate, *D* doubtful, *I* inadequate

TABLE 4

Item generation and reduction – PROM development

	BCL-E-SEI	FL QA-I	FL QA-LS	LL IS	LL IS 2.0	LSI DS-A	LSI DS H&N	LS IDS-L	Lymph-ICF	Lymph-ICF-UL	Lymph-ICF-LL	LYM-QOL	Ly-QLI	PBI-L	(S)LQ-OLI	UL-L-27	ULL-QoL
Item generation																	
Patient interviews																X	
Patient survey		X					X	X	X		X			X			X
Expert opinions						X	X	X							X^a		
Literature review						X	X	X	X		X						X
Based on other PROM		X	X		X		X	X		X			X^a		X^a		
Author's experience				X^a	X												
Item reduction																	
Endorsement frequency						X											
Item redundancy			X			X						X^a					X
Expert opinion		X	X										X^a	X			X
Missing data																	
Factor analysis			X^b										X^a			X	

PROM patient-reported outcome measure

^{a}Limited information on development

^{b}Factor analysis was not conclusive