ACTA ONCOLOGICA
2021, VOL. 60, NO. 3, 379–391
https://doi.org/10.1080/0284186X.2020.1862422

Taylor & Francis
Taylor & Francis Group

ORIGINAL ARTICLE

Self-reported questionnaires for lymphoedema: a systematic review of measurement properties using COSMIN framework

Vincent Singh Paramanandam[a,b], Mi-Joung Lee[a], Sharon L. Kilbreath[a] and Elizabeth S. Dylke[a]

[a]Faculty of Medicine and Health, Sydney School of Health Sciences, The University of Sydney, Sydney, Australia; [b]Physiotherapy Department, Tata Memorial Hospital, Mumbai, India

ABSTRACT

Introduction: Lymphoedema may develop as a result of numerous genetic and traumatic causes; however, treatment for cancer is the most common cause of its development in more economically developed nations. This systematic review critically appraised, compared and summarised the measurement properties of lymphoedema-specific self-reported questionnaires (SRQs) measuring various patient-reported outcomes including quality of life (QOL), function, morbidity, and symptoms.

Methods: Seven databases were searched to identify studies of the measurement properties of SRQs. Two review teams independently evaluated the quality of the individual studies using the risk of bias tool from the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN). Measurement properties of the SRQs presented in the studies were then rated. Study level ratings were summarised for an SRQ if they were reported in multiple studies, and their overall quality of the evidence were then graded.

Results: Forty articles, reporting on 19 SRQs were identified from 8615 records. The focus of the 19 SRQs included eight on QOL, four on symptoms, two on function, and two on impairment. The other three SRQs were on illness perception, self-efficacy, and patient-relevant treatment benefit, respectively. Eight and three SRQs were upper limb and lower limb-specific, respectively, whereas seven questionnaires were for both upper and lower limb lymphoedema. One SRQ was developed for head and neck lymphoedema. According to the COSMIN framework, none of the SRQs reviewed had sufficient evidence to support all nine measurement properties. In lower limb questionnaires, the LYMQOL-leg has sufficient content, structural, and construct validity as well as internal consistency and reliability. For upper limb lymphoedema questionnaires, the Lymph-ICF-UL had sufficient content and construct validity as well as reliability.

Conclusion: LYMQOL-leg SRQ is recommended with confidence for evaluation of QOL of people with lower limb lymphoedema while the Lymph-ICF-UL is recommended for evaluation of the QOL of the breast cancer-related lymphoedema with some confidence. In view of the high level of the indeterminate ratings of the measurement properties of the existing SRQs, further research is desirable.

ARTICLE HISTORY
Received 26 August 2020
Accepted 5 December 2020

KEYWORDS
Lymphoedema; self-reported questionnaire; patient-reported outcome; COSMIN

Background

Lymphoedema is a progressive and disabling condition, characterised by an accumulation of protein-rich fluid in the interstitial spaces of the affected body part. This condition is broadly classified into two categories: (1) primary lymphoedema caused by congenital or hereditary abnormalities of the lymphatic system and (2) secondary lymphoedema caused by cancer treatments, infections, or trauma damaging the lymphatic system. The treatment for cancer is the most common cause of secondary lymphoedema in developed countries [1]. Lymphoedema is a particularly feared consequence of the treatment for cancer as, if not well managed, it may impact negatively on the physical and psychosocial functions as well as the quality of life (QOL) of the people with lymphoedema [2].

Evaluation of lymphoedema should include self-reported questionnaires (SRQs) to capture patient-reported outcomes [3,4], such as symptoms, physical and psychosocial functions as well as QOL. These patient-reported outcomes, apart from objective lymphoedema measures, are increasingly recognised as an integral part of patient-centred care [5,6]. However, to evaluate patient-reported outcomes accurately, the SRQs must have good measurement properties, which include domains like validity, reliability, and responsiveness [5–7]. Validity is the extent to which an SRQ measures the outcome it intends to measure. Reliability is the extent to which the SRQ is free from measurement error. Responsiveness is the ability of an SRQ to detect change over time in the outcome it measures [8]. There are several lymphoedema-specific SRQs currently available to evaluate similar patient-reported outcomes. To inform the selection of

CONTACT Elizabeth S. Dylke Elizabeth.dylke@sydney.edu.au Sydney School of Health Sciences, Susan Wakil Health Building, Western Ave, Camperdown, NSW 2006 Australia

Supplemental data for this article can be accessed here.

an SRQ with good measurement properties, a comprehensive systematic review that evaluates the measurement properties of all lymphoedema-specific SRQs is required [7].

A few systematic reviews with limited scope have evaluated measurement properties of lymphoedema-specific SRQs [9–12]. For example, a recent systematic review reported on the number of domains and symptoms included in SRQs that evaluated QOL of breast cancer-related lymphoedema (BCRL) [10]. However, the review was limited to BCRL QOL SRQs and did not report on the measurement properties. Three other systematic reviews examined studies that had reported the utilisation of the SRQs in the evaluation of QOL of BCRL and cancer-related lower limb lymphoedema [9,11,12]. These reviews reported on the commonly used SRQs, noting the lack of consensus regarding which SRQ to use. To achieve the consensus of use of SRQs, they argued for the need of a comprehensive systematic review of the measurement properties of the lymphoedema specific SRQs using a standard method such as the Consensus-based Standards for the Selection of Health Measurement Instruments (COSMIN) [9]. The current comprehensive systematic review was, therefore, conducted to critically appraise, compare and summarise the measurement properties of lymphoedema-specific SRQs measuring various patient-reported outcomes (constructs) such as QOL, function, morbidity, and symptoms.

Methods

This systematic review was registered prospectively with the International Prospective Register of Systematic Reviews (PROSPERO reference no: CRD42017064645) and conducted based on the COSMIN guideline for systematic reviews [7,13–15].

Database search

The following electronic databases were searched from their date of inception: MEDLINE, Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, AMED, EMBASE, Scopus, and PsycINFO initially in May 2017, with updated searches in March 2019 and April 2020 using the comprehensive search strategy. The search was not restricted for any criteria such as English language or sex, although only studies published in English language were included in this review. The search strategy was developed using keywords from patient-reported outcomes of interest (QOL, function and symptom experience), target population (lymphoedema), and measurement properties identified in COSMIN taxonomy (Supplementary material 1) [8,16]. This search strategy was refined based on the results of a pilot search, and the documentation of the database search for a database (Medline) is given in Supplementary material 2.

Study selection

After removing duplicates, articles were screened by two teams of reviewers (team one: VP and SK; team two: ED and MJ) independently from each other for eligibility using the selection criteria. The titles and abstracts were first screened for eligibility, followed by full-text articles that were not rejected in the initial phase. Any discrepancies in the study selection between reviewers were resolved through discussion. Endnote (Clarivate™, Philadelphia, PA, USA) and Covidence (Melbourne, Australia) were both used for study review and selection.

Selection criteria

Studies were included for this review if they: reported on the development of lymphoedema-specific SRQs; evaluated the measurement properties of lymphoedema-specific SRQs identified within the COSMIN taxonomy (Supplementary material 1) [8]; were lymphoedema-specific SRQs that evaluated any one patient-reported outcome including the QOL, function, morbidity, and symptom experience; and included participants aged 18 or above with primary or secondary lymphoedema, and written in English. Articles were excluded if they were: reviews; studies that used SRQs only as an outcome measurement tool; studies in which only clinician or proxy-reported questionnaires were used; SRQs developed for screening or diagnostic purpose only; studies that included participants with oedema other than lymphoedema; or were not published in English. However, studies that included participants with and without lymphoedema to evaluate the known group validity were included.

Data extraction

Data extraction forms were based on COSMIN Methodology for Systematic Reviews of Patient-Reported Outcome Measures (PROMs) user manual [17]. Two teams of reviewers (team one: VP and SK; team two: ED and MJ) independently extracted the data from each article. Discrepancies in extracted data were resolved through discussion. Data extraction included details related to: (1) development and testing of the SRQs including the patient-reported outcomes and population, recall period, subscales, response options, scoring details, target language, and the completion time; (2) participants' details, details of their lymphoedema and information related to the SRQs administration; and (3) the design of the measurement property studies, the requirements of statistical tests performed and the results for each assessed measurement property [8].

Risk of bias assessment

Two teams of reviewers (team one: VP and SK; team two: ED and MJ) independently assessed risk of bias for each measurement property reported in the study using the COSMIN Risk of Bias checklist [14]. Any discrepancies in the assessment between the teams of reviewers were resolved through consensus. Each measurement property was rated as 'very good', 'adequate', 'doubtful', or 'inadequate' based on standards specific to that measurement property. These standards include requirements for both the design and statistical tests

used to assess the measurement property. The lowest rating from the several standards that comprised the measurement property was its overall rating [14]. Each measurement property included an item which asked if there were any other flaws; no additional flaws were identified and so these were uniformly scored 'very good'. In addition, for evaluating 'test–retest reliability', a time interval between tests of one to two weeks was determined by the team to receive a rating of 'very good'.

Rating of measurement property results

Each of the measurement properties reported in a study were rated using the updated criteria for good measurement properties by two independent teams of reviewers (team one: VP and SK; team two: ED and MJ) [7]. Disagreement was resolved through discussion. A measurement property was rated as 'sufficient (+)' if it met the criteria, 'insufficient (–)' if it failed to meet the criteria, or 'indeterminate (?)' if there was lack of adequate information to determine whether it was sufficient or insufficient [18,19].

Summarising the data

When more than one study reported on the same measurement property for an SRQ, the results from all studies were qualitatively summarised. Using the COSMIN criteria for good measurement properties (V2, 2016 [19]), an overall rating of 'sufficient', 'insufficient', or 'indeterminate' was determined for each measurement property for each SRQ [7,18,19]. The results of the development and content validity studies were also qualitatively summarised, with the overall rating for these areas also determined for each SRQ [15].

Grading the quality of evidence

The overall quality of the evidence was graded as 'high', 'moderate', 'low', or 'very low' using the modified Grading of Recommendation, Assessment, Development, and Evaluation (GRADE) criteria [7]. The quality of the evidence of each reported measurement property was initially graded as 'high' and the level was downgraded step-wise based on following factors: (1) the methodological quality of the studies (risk of bias); (2) imprecision due to small sample size; and (3) indirectness of the result (i.e., evidence from the studies conducted at different setting and different populations than the interest of the review). For ease of interpretation, the quality of the evidence was dichotomised to good and poor-quality evidence; good-quality evidence denotes 'moderate' to 'high' gradings and poor-quality evidence denotes 'low' to 'very low' gradings.

Results

Study selection

The electronic databases search identified 8573 articles and the scoping of the reference lists of the relevant studies and other reviews identified 42 articles. After removing the duplicates and screening the titles and abstracts, 96 articles were selected for full-text review. Out of these 96 full-text articles, 40 articles, reporting on 19 different SRQs (Supplementary material 3), were selected for final review (Figure 1). Articles that were excluded at full-article review are listed in Supplementary material 4, with the reason why they were excluded.

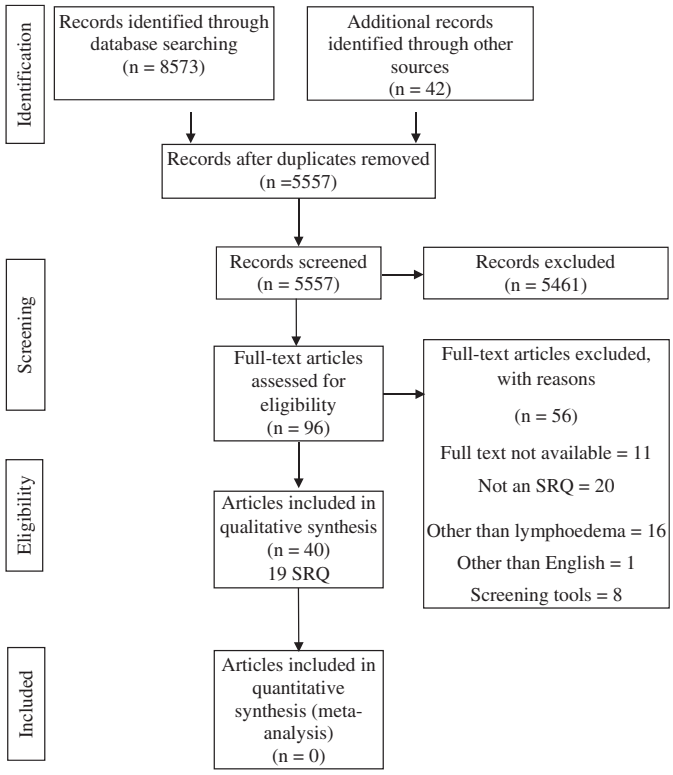

Figure 1. PRISMA flowchart of study selection.

Characteristics of SRQs

Included SRQs were developed for measuring different patient-reported outcomes (Table 1). FLQA-LS, ULLQoL, LyQLI, SLQOLI, LYMQOL-arm, LYMQOL-leg, FLQA-L, and ULL 27 SRQs were for the measurement of the QOL [20–34]. LSIDS-LL, LISIDS-A, BCLE-SEI, and LSIDS-H&N SRQs were for the measurement of the lymphoedema symptoms [35–39]. Lymph-ICF-LL and Lymph-ICF-UL SRQs were for the evaluation of the Functioning, Disability and Health based on The International Classification of Functioning, Disability and Health (ICF) [40–49]. LLIS Version 1 and 2 SRQs were for the evaluation of impairments [50–55]. IPQ-R BCRL, EBS, and PBI-L SRQs were for the evaluation of illness perception, self-efficacy, and patient-relevant treatment benefit, respectively [56–59].

Included SRQs were developed either for a mixed population or a specific patient population. The questionnaires developed for a mixed population, including primary and secondary lymphoedema, were the FLQA-LS, LyQLI, SLQOLI, LYMQOL-leg, FLQA-L, LSIDS-LL, Lymph-ICF-LL, LLIS-V2, LLIS-V1, and PBI-L [20,22–24,26,27,29–32,35,46–55,58,59]. The questionnaires developed exclusively for secondary

Table 1. Characteristics of the included SRQs.

SRQ [Articles]	Construct(s)	Target population	(Sub)scale(s) (number of items)	Response option (range of scores) [scoring interpretation]	Conceptual model (completion time in minutes) [recall period in weeks]	Source language (validated in any other languages)
FLQA-LS [20]	QOL	PL, SL, lip, and lip-LE	Physical impairments (6), daily life (6), social life (6), mental condition (6), therapy of lymphatic disorder (6), global assessment (3)	5-point Likert scale (1–5) [1 = best HRQOL; 5 = worst HRQOL]; VAS for global assessment (0–10) [0 = worst HRQOL; 10 = best HRQOL]	NR (NR) [1]	German (English[a])
ULLQoL [21]	QOL	BCRL	Physical wellbeing (9) and emotional wellbeing (5)	5-point Likert scale (0–56) [0 = best HRQOL; 56 = worst HRQOL]	NR (4.5) [2]	English
LyQLI [22,23]	QOL	PL and SL	Physical (12), psychosocial (16) and practical (13) four additional questions	4-point Likert scale (0–3) [0 = best HRQOL; 3 = worst HRQOL], yes or no for one question, four points Likert for three questions	NR (6) [4]	English[a] (Swedish)
SLQOLI [24]	QOL	PL and SL	Physical (51), emotional (48), social (30), practical (54), general HRQOL items (2), and open-ended questions (2)	4-point Likert scale (0–3) [0 = best HRQOL; 3 = worst HRQOL], VAS for two QOL question (0–10) [0 = poor HRQOL; 10 = excellent HRQOL] and two open ended question	NR (30) [4]	English[a] (Swedish)
LYMQOL-arm [25–28]	QOL	LE	Functional (10), appearance (5), symptoms (6), mood (6), and overall QOL (1)	4-point Likert scale (1–4) [1 = best HRQOL; 4 = worst HRQOL], VAS for one general QOL question (0–10) [0 = poor HRQOL; 10 = excellent HRQOL]	NR (NR) [1[b]]	English (Swedish and Turkish)
LYMQOL-leg [26,27, 29–31]	QOL	LE	Functional (8), appearance (7), symptoms (5), mood (6), and overall QOL (1)	4-point Likert scale (1–4) [1 = best HRQOL; 4 = worst HRQOL], VAS for one general QOL question (0–10) [0 = poor HRQOL; 10 = excellent HRQOL]	NR (NR) [1[b]]	English (Turkish, Swedish and Dutch)
FLQA-L [32]	QOL	PL and SL	Physical complaints (17), everyday life (14), social life (7), emotional wellbeing (26), treatment (8), satisfaction (13), and household/work (7)	5-point Likert scale (1–5) [1 = best HRQOL; 5 = worst HRQOL]	NR (NR) [1[c]]	German
ULL 27 [33,34]	QOL	BCRL	Physical dimension (15), psychological dimension (7), and social dimension (5)	5-point Likert scale (0–100) [0 = worst HRQOL; 100 = best HRQOL]	NR (11) [4]	French (Dutch)
LSIDS-LL [35]	Symptom experience	PL and SL	Activity (5), soft tissue sensation (4), pain (4), resources (2), biobehavioural (7), neurological sensation (3), function (3), and sexuality (3)	Yes or no for symptoms occurrence subscale (0–36) [0 = best; 36 = worst], 4-point Likert scale for Intensity and Distress scale (0–10) [0 = best; 10 = worst]	The Lenz Theory of Unpleasant Symptoms (NR) [1]	English
LSIDS-A [36]	Symptom experience	BCRL	Activity (3), soft tissue sensation (4), resources (2), biobehavioural (9), neurological sensation (7), function (2) and sexuality (3)	Yes or no for symptoms occurrence subscale (0–30) [0 = best; 30 = worst], 4-point Likert scale for Intensity and Distress scale (0–10) [0 = best; 10 = worst]	The Lenz Theory of Unpleasant Symptoms (12) [1]	English
BCLE-SEI [37,38]	Symptom experience	BCRL	Symptom occurrence (28) and symptom distress (28)	Yes or no for symptoms occurrence subscale (0–24) [0 = best; 24 = worst], five-point Likert scale for Intensity and Distress scale (0–120) [0 = best; 128 = worst]	NR (NR) [4]	English (Chinese)
LSIDs-H&N [39]	1. Symptom burden; 2. Functional loss	HNLE	Altered sensation (14), psychosocial issues (17), functional impairments and symptom burden (27) and perception of tissue swelling (9)	Yes or no for symptom presence, VAS for symptom severity and symptom distress (1–10) [1 = slight; 10 = severe]	NR (10) [1]	English
Lymph-ICF-UL [40–45]	Function	BCRL	Physical (7), mental (4), household (4), mobility (8), and life and social activities (6)	VAS (0–100) [0 = best score; 100 = worst score]	ICF (5) [2]	Dutch (French, Turkish and Danish)

(*continued*)

Table 1. Continued.

SRQ [Articles]	Construct(s)	Target population	(Sub)scale(s) (number of items)	Response option (range of scores) [scoring interpretation]	Conceptual model (completion time in minutes) [recall period in weeks]	Source language (validated in any other languages)
Lymph-ICF-LL [46–49]	Function	PL and SL	Physical function (6), mental function (6), general tasks/household activities (3) and mobility (7), life domains/social life (13)	11-point VAS (0–100) [0 = best score; 100 = worst score]	ICF (5) [2]	Dutch (Brazilian Portuguese, Turkish and Chinese)
LLIS-V2 [50–52]	Impairment	PL and SL	Physical concerns (6), psychosocial concerns (6), functional concerns (5), and infection occurrence (1)	5-point Likert scale (0–72) [0 = best; 72 = worst]	NR (NR) [1]	English (Turkish and Arabic)
LLIS-V1 [53–55]	Impairment	PL and SL	Physical concerns (8), psychosocial concerns (4), and functional concerns (6)	5-point Likert scale (18–90) [18 = best; 90 = worst]	NR (NR) [2]	English (Persian and Danish)
IPQ-R BCRL [56]	Illness perception	At-risk of BCRL	Identity subscale (8), causal subscale (22), 38 items (measuring: timeline, consequences, personal control and treatment control, illness coherence, timeline cyclical, and emotional representation components)	Yes or no for identity subscale (0–8) [0 = no symptom identified with BCRL; 8 = highest number of symptoms identified with BCRL], five-point Likert scale for causal subscale (1–22) [22 = higher degree of approval that the item may result in BCRL] and five-point Likert scale for the remaining 38 items (1–38) [higher the score, stronger the belief was held in a specific illness component]	The Common-Sense Model of Self-Regulation (15) [present]	Chinese
EBS [57]	Self-efficacy for overcoming LE-related exercise barriers	CRLE	Exercise barriers self-efficacy (5)	11-point VAS (0–100%) [0 = not at all confident; 100% = extremely confident]	NR (NR) [present]	English
PBI-L [58,59]	Patient-relevant treatment benefit	LE and lip	Normal everyday-life and psychological wellbeing (14) and physical wellbeing and capability (9)	5-point Likert scale (0–4) [0 = no benefit; 4 = maximum benefit]	NR (NR) [[b]]	German (English[a])

SRQ: self-reported questionnaire; PL: primary lymphoedema; SL: secondary lymphoedema; Lip: lipedema; LipLE: lipo-lymphoedema; VAS: visual analogue scale; HRQOL: health-related quality of life; NR: not reported; BCRL: breast cancer-related lymphoedema; QOL: quality of life; LE: lymphoedema; HNLE: head and neck lymphoedema; ICF: International Classification of Functioning, Disability, and Health; CRLE: cancer-related lymphoedema; FLQA-LS: Freiburg Life Quality Assessment for patients with lymphedema short; ULLQoL: Upper Limb Lymphedema Quality of Life Questionnaire; LyQLI: Lymphedema Quality of Life Inventory; SLQOLI: Swedish version of Lymphedema Quality of Life Inventory; LYMQOL-arm: quality of life measure for limb lymphoedema – arm; LYMQOL-leg: quality of life measure for limb lymphoedema – leg; FLQA-L: Freiburg Life Quality Assessment for patients with lymphedema; ULL 27: upper limb lymphoedema 27; LSIDS-LL: Lymphedema Symptom Intensity and Distress Survey-Lower limb; LSIDS-A: Lymphedema Symptom Intensity and Distress Survey-Arm; BCLE-SEI: Breast Cancer and Lymphedema Symptom Experience Index; LSIDs-H&N: Lymphedema Symptom Intensity and Distress Survey-Head and Neck; Lymph-ICF-UL: Lymphoedema Functioning, Disability and Health Questionnaire for Upper Limb; Lymph-ICF-LL: Lymphoedema Functioning, Disability and Health Questionnaire for Lower Limb; LLIS V2: Lymphedema Life Impact Scale V2; LLIS V1: Lymphedema Life Impact Scale V1; IPQ-R-BCRL: revised illness perception questionnaire BCRL; EBS: exercise barriers self-efficacy; PBI-L: patient benefit index in lymphedema and lipedema.

[a]Validation study is not published.

[b]Varying recall period.

[c]For one or more items the recall period is 'present'.

lymphoedema were the ULLQoL, LYMQOL-arm, ULL 27, LSIDS-A, BCLE-SEI, LSIDs-H&N, Lymph-ICF-UL, IPQ-R BCRL, and EBS [21,26–28,34,36–42,45,56,57]. These questionnaires were developed for cancer-related lymphoedema of the upper limb, such as BCRL, except the LSIDs-H&N and Exercise barriers self-efficacy questionnaires, which were for head and neck lymphoedema and any cancer-related lymphoedema, respectively [39,57]. Although IPQ-R BCRL included only women who are at-risk of BCRL in the validation study, this SRQ was included in this review (Table 1).

Other characteristics of the SRQs, such as recall period, number of subscales, number of items, response options, and language availability varied among the reviewed SRQs. The recall period ranged from the present [56,57] to the past four weeks [22–24,33,34]. However, the recall period for three SRQs either varied or was not explicitly presented for some items [25–31,58,59]. Except for the Exercise Barrier Self-Efficacy SRQ [57], which had no subscales, other SRQs had between two [21,37,38] and eight subscales [35]. The number of items in the SRQs ranged from 5 to 188 items [24,57]. Combinations of scales, such as the Likert, visual analogue scale, or dichotomous (yes or no), were used in 10 SRQs [20–23,26–28,30,31,35–39,55]. Likert type scales only were used in six SRQs [32–34,50–55,58,59], and visual analogue scales only were used in three SRQs [40–49,57]. With inclusion of the translated questionnaires, 10 SRQs were in English [21,26,35,36,38,39,52,55], and 17 SRQs were in other languages including: German [20,32,59]; Dutch [29,33,40,42,45,49]; Turkish [25,28,30,31,44,47,51,53,58]; Chinese [37,46,56]; Brazilian Portuguese [48]; French [34,41]; Danish [43]; Arabic [50]; Persian [54]; and Swedish [27]. Other characteristics and available translations are presented in Table 1.

Characteristics of participants

Overall, 5833 participants were included in the 40 studies that evaluated the measurement properties of the 19 SRQs (Table 2) [20–59]. Participants were predominantly women (89.9%) with a mean age ranging from 48 to 64 years. Sample size varied among the included studies from 10 [48] to 446 [54]. Eighteen studies were conducted exclusively on women [25,28,33,34,36–38,40–46,48,51,56,57]. The severity and duration of lymphoedema were not reported in all studies, and those reported did not report them uniformly (Table 2). Other details regarding the instrument administration are presented in Table 2.

Summary of findings

Seven of the nine measurement properties identified in the COSMIN taxonomy were rated and graded for the SRQs included in this review (Supplementary material 3 and Table 3). Although eight SRQs were translated into different languages, none of them could be rated for cross-cultural adaptation because they did not assess measurement invariance [24,25,27–31,33,37,41,43,44,46–48,50,51,53,54,58]. Similarly, despite two of the included SRQs reporting on criterion validity, this finding was not rated due to lack of a gold standard SRQ as a criterion with which they could be compared [26,37].

From the 19 SRQs reviewed, the ULLQoL, LYMQOL-arm, LYMQOL-leg, LSIDs-H&N, Lymph-ICF-UL, Lymph-ICF-LL, LLIS Version 1 SRQs were rated as sufficient for content validity with good-quality evidence (Table 3) [21,27,28,31,36,39,45,49,52,55,59]. The LYMQOL-leg SRQ was also rated sufficient for structural validity, internal consistency, and hypothesis testing with good-quality evidence [26,27,29–31]. Only content validity was reported for the LSIDs-H&N SRQ. The LYMQOL-arm, Lymph-ICF-LL, and LLIS Version 1 were rated insufficient for structural validity based on good-quality evidence. The structural validity of the ULLQoL was not graded due to indeterminate rating, and the structural validity for the Lymph-ICF-UL was not reported. Although high Cronbach's alphas for internal consistency were reported for 13 SRQs, internal consistencies were rated indeterminate due to lack of evidence for their structural validity [21,25–28,40–49,53–55]. The hypothesis testing for the ULLQoL and Lymph-ICF-UL were rated sufficient based on good-quality evidence. The responsiveness for the ULLQoL and the reliability for the Lymph-ICF-UL were rated sufficient based on good-quality evidence. Quality grading of only IPQ-R SRQ [56] was downgraded since the developers included only at-risk women in the one available study rather than women with BCRL. Overall, only 3% of the reported measurement properties were rated as insufficient based on good-quality evidence (Table 3).

Discussion

Nineteen lymphoedema-specific SRQs were identified, and their measurement properties critically appraised and summarised. According to the COSMIN framework, none of the SRQs was rated as sufficient based on good-quality evidence for all nine measurement properties. However, LYMQOL-leg and Lymph-ICF-UL SRQs were rated as sufficient with respect to content validity, reliability, and construct validity based on good-quality evidence [26,27,29–31,40–45]. Each of the 17 other SRQs had one or more measurement properties that were rated as sufficient and supported by good quality evidence. Overall, only 3% of the reported measurement properties were rated as insufficient based on good-quality evidence and therefore not acceptable. However, in all SRQs, the criteria for at least one or more measurement properties were found to be indeterminate (33%) or not reported on at all (33%). Reasons why criteria were indeterminate were related to lacking details in the reporting of the methods, whereas for others, it may reflect lack of knowledge in the developers on the steps involved in establishing the acceptability of measurement properties.

The rating of 'sufficient' for a measurement property is given when studies have met the criteria outlined by COSMIN's updated criteria for good measurement properties [7]. However, the extent to which risk of bias is minimised in these studies determines the quality of evidence. Good-

Table 2. Characteristics of the included study populations.

Articles (SRQ)	*N* (% female)	Age – mean (SD) in years	Type in %	Affected body part	Duration: mean (SD), months	Severity ISL stages in %	Country (language)
Augustin 2018 [20] (FLQA-LS)	348 (90.8)	57.3 (24–89)[a]	PL=22.3 SL=44.9 Lip=9.6 LipLE=23.3	UL and LL	NR	NR	Germany (German)
Williams 2018 [21] (ULLQoL)	103 (97)	60.5 (13)	SL (BCRL=99; NHL=1)	UL	37 (12)	NR	United Kingdom (English)
Klernas 2015 [23] (LyQLI)	126 (87)	62 (19–92)	SL=74 PL=26	UL and LL	84 (0–840)[b]	NR	Sweden (Swedish)
Klernäs 2018 [22] (LyQLI)							
RP	18 (83)	61 (46–76)[b]	PL=22 SL=78	UL and LL	30 (6–684)[b]	NR	
LipoSuc	50 (90)	55 (23–75)[b]	PL=30 SL=70	UL and LL	144 (24–792)[b]	NR	Australia, Scotland, and Sweden (Swedish)
LE	129 (86)	62 (19–92)[b]	PL=26 SL=74	UL and LL	84 (6–840)[b]	NR	
Klernas 2010 [24] (SLQOLI)	76 (88)	63 (25–84)[b]	SL=74 PL=26	LL and UL	84 (0–840)[b]	NR	Sweden (Swedish)
Keeley 2010 [26] (LYMQOL-arm and leg)	209 (78.7)	58 (16.4)	NR	UL and LL	NR	NR	United Kingdom (English)
Borman 2018 [25] (LYMQOL-Arm)	135 (100)	51.8 (9.8)	SL (BCRL)	UL	21.1 (38.7)	1 = 29 2 = 62 3 = 1	Turkey (Turkish)
Karayurt 2019 [28] (LYMQOL-Arm)	109 (100)	55.69 (9.33)	SL (BCRL)	UL	39.36 (34.92)	NR	Turkey (Turkish)
Wedin 2020 [27] (LYMQOL-arm)	52 (98)	62 (34–75)[b]	NR	UL	2 (0.1–20)[b]	NR	Sweden (Swedish)
van de Pas 2016 [29] (LYMQOL-leg)	76 (68.7)	59.6 (15.6)	PL=14.9 SL=53.7Unknown=31.3	LL	NR	NR	Netherlands (Dutch)
Bakar 2019 [30] (LYMQOL-leg)	119 (72.3)	49.83 (14.4)	PL=45.4 SL=54.6	LL	84 (75)	2 = 30 3 = 70	Turkey (Turkish)
Wedin 2020 [27] (LYMQOL-arm and leg)	50 (94)	61 (34–82)[b]	NR	LL	1.5 (2–33)[b]	NR	Sweden (Swedish)
Borman 2020 [31] (LYMQOL-leg)	138 (86.2)	52.01 (14.7)	PL=47.1 SL=52.9	LL	95.6 (108.6)	1 = 14 2 = 75 3 = 11	Turkey (Turkish)
Augustin 2005 [32] (FLQA-LS)	177 (84)	49.5 (18–77)[a]	PL=63.4 SL=37.6	NR	NR	NR	Germany (German)
Launois 2002 [34] (ULL 27)	301 (100)	61.6 (1.2)	SL (BCRL)	UL	NR	1 = 13 2 = 20 3 = 29 4 = 38	France (French)
Viehoff 2007 [33] (ULL 27)	84 (100)	59 (11.8)	SL (BCRL)	UL	35.5 (45.1)	NR	Netherlands (Dutch)
Ridner 2015 [36] (LSIDS-A)	236 (100)	58.9 (11)	SL (BCRL)	UL	29.3 (0–73.8)[b]	2 = 85	USA (English)
Fu 2016 [38] (BCLE- SEI)	355 (100)	21–39 *y* = 10.4%; 40–59 *y* = 53.0%; 60–80 *y* = 36.6%	SL (BCRL)	UL	NR	NR	USA (English)
Shi et al. 2016 [37] (BCLE-SEI)	219 (100)	57.6 (9.6)	SL (BCRL)	UL	NR	NR	China (Chinese)
Ridner 2018 [36] (LSIDS-LL)	388 (86.6)	52 (13.4)	PL=51.4 SL=48.6	LL	67.8 (0.0–657.6)[b]	NR	USA (English)
Deng 2012 [39] (LSIDs-H&N)	30 (23.30)	60.4 (40.4–81.4)[b]	SL	HN	NR	0 = 37 1 = 60 2 = 3	USA (English)
Devoogdt 2011 [45] (Lymph-ICF-UL)	90 (100)	OB=61.2 (10.0); SU=56.7 (9.3); WOL-58.3 (11.9)	SL	UL	OB=41 SU=19	NR	Belgium (Dutch)
Kostanoglu 2016 [44] (Lymph-ICF-UL)	30 (100)	53.8 (5.8)	SL (BCRL)	UL	18.2 (7.7)	NR	Turkey (Turkish)
Grarup 2019 [43] (Lymph-ICF-UL)	77 (100)	61 (12.4)	SL	UL	15.5 (58)[c]	NR	Denmark (Danish)

(*continued*)

Table 2. Continued.

Articles (SRQ)	*N* (% female)	Age – mean (SD) in years	Type in %	Affected body part	Duration: mean (SD), months	Severity ISL stages in %	Country (language)
De Vrieze 2019 [42] (Lymph-ICF-UL-revision)	56 (100)	62 (10)	SL (BCRL)	UL	34.5 (13.5 and 79.5)[d]	1 = 18 2*a*=59 2*b* = 23	Belgium (Dutch)
De Vrieze 2020 [40] (Lymph-ICF-UL)	50 (100)	64 (11)	SL (BCRL)	UL	78 (30 and 177)[d]	1 = 0 2*a*=38 2*b* = 62	Belgium (French)
De Vrieze 2020 [41] (Lymph-ICF-UL-responsiveness)	95 (100)	62 (10)	SL (BCRL)	UL	53 (42.5)	1 = 13 2*a*=59 2*b* = 28	Belgium (Dutch)
Devoogdt 2014 [49] (Lymph-ICF-LL)	30 (80)	50.9 (14.9)	SL	LL	NR	NR	Belgium and Netherland (Dutch)
Ferreira 2016 [48] (Lymph-ICF-LL)	10 (100)	54 (12.3)	SL (cancer)	LL	31.2 (59.3)	1 = 20 2 = 40 3 = 30 4 = 0	Brazil (Brazilian Portuguese)
Kostanoğlu 2017 [47] (Lymph-ICF-LL)	50 (72)	50.3 (12.9)	PL=64 SL=36	LL	92.9 (34.9)	1 = 14 2 = 86	Turkey (Turkish)
Wang 2018 [46] (Lymph-ICF-LL)	170 (100)	54.4 (10.7)	SL (GLE)	LL	NR	NR	Taiwan (Chinese)
Weiss 2018 [52] (LLIS-V2)	84 (82)	UL=59.9 (11.9) LL=59.6 (14.1)	PL=12 SL=88	UL and LL	NR	NR	USA (English)
Orhan 2019 [51] (LLIS-V2)	113 (100)	56.5 (10.2)	SL (BCRL)	UL	NR	NR	Turkey (Turkish)
Abu Sharour 2020 [50] (LLIS-V2)	90 (NR)	44.1 (1.1)	SL (BCRL)	UL	NR	NR	Jordan (Arabic)
Weiss 2015 [55] (LLIS-V1)	102 (85.3)	58.7 (14.1)	PL=12.7 SL=87.3	UL and LL	NR	NR	USA (English)
Haghighat 2018 [54] (LLIS-V1)	446 (NR)	LE=53.3 (10.9) WOL=51.5 (10.5)	SL (BCRL)	UL	NR	NR	Iran (Persian)
Degirmenci 2019 [53] (LLIS-V1)	106 (97)	53.6 (±11.8)	PL=18.6 SL=81.4	UL and LL	UL=24 (1–396) LL=54 (1–384)[b]	UL: 1 = 13 2 = 87 LL: 1 = 4 2 = 82 3 = 15	Turkey (Turkish)
Huang 2019 [56] (IPQ-R-BCRL)	204 (100)	48.0 (±9.6)	At-risk of BCRL	UL	NA	NR	China (Chinese)
Buchan 2015 [57] (EBS)	38 (100)	56.3 (1.6)	SL (cancer)	UL	NR	NR	Australia (English)
Blome 2014 [59] (PBI-L)	383 (92.8)	CS=57.4 (14.3) LS=60.2 (13.9)	PL=12.9, SL=43.5, PL&SL=1 Lip or LipLE=33.6	LL and UL	NR	NR	CS – Germany (German); LS – USA (English)
Duygu 2020 [58] (PBI-L)	81 (63)	47.66	PL=43.2, SL=49.4, PL&SL=3.7 Lip or LipLE=3.7	LL and UL	NR	1 = 27 2 = 54 3 = 189	Turkey (Turkish)

SRQ: self-reported questionnaires; *N*: sample size; SD: standard deviation; ISL: International Society of Lymphology; PL: primary lymphoedema; SL: secondary lymphoedema; Lip: lipedema; LipLE: lipo-lymphoedema; UL: upper limb; LL: lower limb; HN: head and neck; NR: not reported; NA: not applicable; BCRL: breast cancer-related lymphoedema; NHL: non-Hodgkin's lymphoma; RP: rehabilitation program; LipoSuc: liposuction; LE: lymphoedema; OB: objective lymphoedema; SU: subjective lymphoedema; WOL: without lymphoedema; CS: cross sectional study; LS: longitudinal study. FLQA-LS: Freiburg Life Quality Assessment for patients with lymphedema short; ULLQoL: Upper Limb Lymphedema Quality of Life Questionnaire; LyQLI: Lymphedema Quality of Life Inventory; SLQOLI: Swedish version of Lymphedema Quality of Life Inventory; LYMQOL-arm: quality of life measure for limb lymphoedema – arm; LYMQOL-leg: quality of life measure for limb lymphoedema – Leg; FLQA-L: Freiburg Life Quality Assessment for patients with lymphedema; ULL 27: upper limb lymphoedema 27; LSIDS-LL: Lymphedema Symptom Intensity and Distress Survey-Lower limb; LSIDS-A: Lymphedema Symptom Intensity and Distress Survey-Arm; BCLE-SEI: Breast Cancer and Lymphedema Symptom Experience Index; LSIDs-H&N: Lymphedema Symptom Intensity and Distress Survey-Head and Neck; Lymph-ICF-UL: Lymphoedema Functioning, Disability and Health Questionnaire for Upper Limb; Lymph-ICF-LL: Lymphoedema Functioning, Disability, and Health Questionnaire for Lower Limb; LLIS V2: Lymphedema Life Impact Scale V2; LLIS V1: Lymphedema Life Impact Scale V1; IPQ-R-BCRL: revised illness perception questionnaire BCRL; EBS: exercise barriers self-efficacy; PBI-L: patient benefit index in lymphedema and lipedema.

[a]Mean and range.

[b]Median and range.

[c]Median and interquartile range.

[d]Median and 25th and 75th percentile.

Table 3. Summary of findings of the measurement properties of the included SRQs.

SRQ	**Content validity**	Structural validity	Internal consistency	Reliability	Measurement error	**Hypotheses testing (construct validity)**	Responsiveness
FLQA-LS	NR	NR	?(NG)[a] $\alpha = 0.79–0.83$	NR	NR	**+(M)**	NR
ULLQoL	**+(M)**	?(NG)	?(NG)[a] $\alpha = 0.77–0.87$	+(V.L.)	NR	**+(M)**	+(L)[d]
LyQLI	?(NG)	?(NG)	?(NG)[a] $\alpha = 0.88–0.92$	**+(M)**	NR	**+(M)**	+(L)[e]
SLQOLI	?(NG)	NR	NR	–(VL)	?(NG)	+(L)	NR
LYMQOL-arm	**+(M)**	–(H)	?(NG)[a] $\alpha = 0.79–0.9$	–(H)	?(NG)	**+(H)**[b]	NR
LYMQOL-leg	**+(M)**	**+(H)**	**+(H)**	**+(H)**	?(NG)	**+(H)**	?(NG)[e]
FLQA-L	?(NG)	NR	?(NG)[a] $\alpha = 0.87–0.95$	+(L)	NR	+(L)[b]	+(VL)[e]
ULL 27	?(NG)	–(H)	?(NG)[a] $\alpha = 0.78–0.93$	**+(M)**	NR	**+(M)**[b]	+(L)[e]
LISIDs-LL	NR	NR	?(NG)[a] $\alpha = 0.76–0.89$	NR	NR	**+(M)**[b]	NR
LSIDS-A	+(L)	NR	?(NG)[a] $\alpha = 0.93–0.94$	?(NG)	NR	**+(M)**	NR
BCLE-SEI	NR	?(NG)	?(NG) $\alpha = 0.92–0.97$	?(NG)	NR	**+(H)**[b]	+(VL)[e]
LSIDs-H&N	**+(M)**	NR	NR	NR	NR	NR	NR
Lymph-ICF-UL	**+(H)**	NR	?(NG)[a] $\alpha = 0.72–0.99$	**+(M)**	+(L)	**+(M)**[b]	+(L)[c,d,e]
Lymph-ICF-LL	**+(M)**	–(M)	?(NG)[a] $\alpha = 0.82–0.97$	**+(M)**	+(VL)	**+(M)**	NR
LLIS Ver.2	±(VL)	NR	?(NG)[a] $\alpha = 0.85–0.95$	+(VL)	NR	+(VL)	+(VL)[e]
LLIS ver. 1	**+(M)**	–(H)	?(NG)[a] $\alpha = 0.84–0.93$	+(VL)	NR	**+(H)**[b]	NR
IPQ-R[BCRL]	?(NG)	+(L)	**+(M)**	NR	NR	NR	NR
EBS	+(L)	+(L)	+(L)	–(VL)	NR	–(VL)	NR
PBI-L	?(NG)	**+(M)**	**+(H)**	NR	NR	**+(M)**	–(L)[c]

SRQ: self-reported questionnaires; +: sufficient rating; –: insufficient rating; ?: indeterminate rating; (): grading of overall quality of evidence based on modified GRADE approach; H: high; M: moderate; L: low; VL: very low; NG: not graded; NR: not reported.
Bold denotes measurement criteria that have sufficient rating and a GRADE of moderate to high.
[a]Criteria for the unidimensionality is not met.
[b]Hypothesis testing including known-group validity.
[c]Responsiveness-comparison with other outcome measurement instruments.
[d]Responsiveness-comparison between subgroups.
[e]Responsiveness-before and after intervention.

quality evidence implies that all steps taken to study that measurement property have used appropriate methods, sample size, and participants [14]. It is the combination of the rating of meeting the required criteria as well as the quality of evidence supporting it that determines whether a measurement property of an SRQ is or is not fully addressed (Table 4). A measurement property is not addressed if it is rated as 'insufficient' and based on good-quality evidence. It is questionable whether further research would address these significant issues. For all other combinations, further evidence through either improved documentation or further research is required to indicate that the measurement property is fully addressed. In the current review, and as highlighted in Table 3, the majority of the measurement properties of most SRQs required further evidence.

Inadequate reporting contributed to a measurement property receiving an indeterminate rating, a common occurrence identified both by the developers of COSMIN framework [14,15] as well in others who have reviewed SRQs using COSMIN methodology [60,61]. This inadequacy may be related to lack of guidelines for reporting of studies that develop or evaluate the measurement properties of SRQs [62,63]. The implementation of guidelines has improved reporting of other types of studies such as the CONSORT guidelines used to report randomised controlled trials [64,65]. The lack of guidelines for reporting on the development of SRQs is currently being addressed, with a guideline currently under development [66]. Meanwhile, evaluation frameworks, such as COSMIN, can be used by developers of an SRQs to report against criteria for each measurement property they have addressed [7,13,14,62,67].

The first step in determining the usability of an SRQ is in establishing its content validity as this step determines that it is relevant, complete and understandable for the intended population [7,68]. Without addressing this measurement property, investigation of other measurement properties is not warranted [15,68]. In this review, only seven SRQs had acceptable content validity: LYMQOL-leg, ULLQoL, LYMQOL-arm, LSIDs-H&N, Lymph-ICF-UL, Lymph-ICF-LL, and LLIS Version 1 SRQs [21,25–31,39–55]. The other SRQs did not meet the criteria for content validity as they did not involve the target population in content development or validation [35,56] or did not describe their methods in enough detail to meet the criteria of COSMIN [20,22–24,26,32–34,36–38,57–59]. One reason why some SRQs may not have addressed the criteria was that the COSMIN criteria were only published in 2018; the majority (65%) of the studies in this review were conducted prior to 2018. Since none of the SRQ was rated insufficient for content validity, further evaluation is warranted of this measurement property in consultation with the content validity criteria may improve the content validity of the existing SRQs.

Another step that was not adhered to in developing the SRQs was related to the assessment of internal consistency. Internal consistency is, in part, conditional on first establishing the acceptability of structural validity [7,69]. Structural validity ensures the score of the SRQ is an adequate reflection of the construct (outcome) and dimensions of the construct (subscales), whereas internal consistency ensures that interrelatedness of the items in the scale or subscales [7]. Of the seven SRQs in which content validity was addressed,

Table 4. Decision matrix for the interpretation of the results.

Decision matrix	Sufficient	Insufficient	Indeterminant
Good-quality evidence (high or moderate grade)	Acceptable, can be used with confidence.	Not acceptable, cannot be used. Further evaluation may not improve the rating.	Not acceptable, since lack of adequate information (due to inadequate reporting) to rate the measurement property as sufficient or insufficient. Grading of the quality of evidence was not done. Further evaluation with good quality studies is needed.
Poor-quality evidence (low or very low grade)	Not acceptable but can be used with caution. Further evaluation with good quality studies is needed.	Not acceptable, cannot be used until further evaluation with good quality studies is done.	

only the LYMQOL-leg was addressed for internal consistency [26,27,29–31]. This was the only SRQ that had undergone confirmatory factor analysis to confirm the subdomains within the SRQ. The other SRQs did not undertake this analysis, although they did report high Cronbach's alphas, i.e., internal consistency, for their SRQs [21,26–28,31,39–53,55]. Establishment of the subdomains using confirmatory factor analysis for these other SRQs would establish both their structural validity and internal consistency.

For SRQs to be useful in clinical practice, it is important to ensure that they are reliable, with low measurement error, and are responsive to changes in the condition for which they were designed. Although test–retest reliability and measurement error are interrelated measurement properties, measurement error was less reported than test–retest reliability. Of the seven SRQs that had the acceptable content validity, only LYMQOL-leg and Lymph-ICF-UL had acceptable reliability whereas both of them did not have acceptable measurement error [26,27,29–31,40–45]. Small sample sizes of the retest were a common problem contributing to poor quality evidence for both test–retest reliability and the measurement error in the included studies [45,49]. Studies with adequate retest sample size are, therefore, required to strengthen the evidence of test–retest reliability and measurement error existing SRQs. In addition, none of the reviewed SRQs had acceptable responsiveness. Lack of acceptable responsiveness of the SRQs was due to the low methodological quality of the studies and low sample size [27,29,31,36,46,49]. Since longitudinal studies are required for the evaluation of the responsiveness, longitudinal observational or intervention trials in which these SRQs are used may be an efficient approach to addressing this measurement property.

Cross-cultural adaptation of existing SRQs assists in comparing outcomes across different cultures [70,71]. This cross-cultural adaptation process consists of two steps: (1) translation of an existing SRQ followed by ascertainment of the translated version's equivalence with the original SRQ and (2) validation of the translated questionnaire [71]. All 11 cross-culturally adapted SRQs followed acceptable translation and equivalence processes and reported various measurement properties that were evaluated during the validation process [24,25,29,30,33,37,43,44,46–48,54]. However, none of these cross-culturally adapted SRQs analysed the measurement invariance, an essential step in validation for cross-cultural validity [14,19]. In brief, measurement invariance compares the responses to items from two cohorts that are similar in respect to the population characteristics other than the language [7,14,19,72]. Omission of this step may be due to a lack of awareness of this critical step. Evaluation of the measurement invariance of the translated SRQs is, therefore, required.

Overall, it was disappointing that so few SRQs could be supported in this review. One issue why this occurred is related to how the score for the measurement property is determined within the COSMIN framework. The score is based solely on the lowest score across all of the criteria underpinning that property, which range from three criteria for criterion validity to eight criteria for reliability. This stringent approach led to harsh ratings for methodological quality that may simply have reflected lack of sufficient documentation around a study design, an issue also identified by others [62,63]. Based on the scoring for COSMIN risk of bias tool, several reviews have concluded that most of the included studies in their reviews were of poor methodological quality, with inconclusive evidence to support any of the reviewed SRQs [60–62,73]. In our current review, this issue was overcome, in part for LYMQOL-leg, and Lymph-ICF-UL and LLIS SRQs, when multiple studies evaluated the measurement properties of the same SRQs [26,27,29–31,40–45]. Additional studies can be undertaken by others to address gaps in the evidence for specific measurement properties. Furthermore, the reporting guidelines currently in development [66] should improve documentation of the steps taken to investigate the measurement property.

In this review, every effort was made to include all the studies that evaluated the measurement properties of lymphoedema specific SRQs. One limitation of our study is that we only included studies that were reported in English. As a consequence, one SRQ in which the evaluation was published in Japanese was excluded [74]. On the other hand, one of the strengths of this review is its comprehensiveness and adherence to the Prisma protocol (Supplementary material 5) [75]. We included all lymphoedema-specific SRQs, unlike previous systematic reviews which restricted their inclusion to either breast or gynaecological cancer related lymphoedema-specific SRQs only [10–12]. Another strength of this review is that we used the COSMIN methodology to conduct the systematic review, including the new COSMIN risk of bias tool, the updated criteria for measurement properties and the modified GRADE approach to grade the overall quality of evidence [7,8,14,15,19]. Use of a modified GRADE approach facilitated the process of recommendation based on the overall quality of evidence synthesised from different studies similar to the systematic reviews of effectiveness trials.

Conclusion

In summary, this systematic review revealed that out of 19 SRQs included in this review, the LYMQOL-leg SRQ could be

recommended with confidence for evaluation of QOL of people with lower limb lymphoedema. Furthermore, the Lymph-ICF-UL can be recommended for evaluation of the QOL of the BCRL with some confidence. In view of the high level of the indeterminate ratings and lack of reporting some measurement properties of all existing SRQs, further validation studies are desirable.

Disclosure statement

No potential conflict of interest was reported by the author(s).

ORCID

Vincent Singh Paramanandam http://orcid.org/0000-0002-9976-9172
Mi-Joung Lee http://orcid.org/0000-0001-6622-3459
Sharon L. Kilbreath http://orcid.org/0000-0002-8024-8641
Elizabeth S. Dylke http://orcid.org/0000-0002-7115-8142

References

[1] Rockson SG, Keeley V, Kilbreath S, et al. Cancer-associated secondary lymphoedema. Nat Rev Dis Primers. 2019;5(1):22.
[2] Fu MR, Ridner SH, Hu SH, et al. Psychosocial impact of lymphedema: a systematic review of literature from 2004 to 2011. Psychooncology. 2013;22(7):1466–1484.
[3] McLaughlin SA, Staley AC, Vicini F, et al. Considerations for clinicians in the diagnosis, prevention, and treatment of breast cancer-related lymphedema: recommendations from a Multidisciplinary Expert ASBrS Panel: Part 1: definitions, assessments, education, and future directions. Ann Surg Oncol. 2017;24(10):2818–2826.
[4] Levenhagen K, Davies C, Perdomo M, et al. Diagnosis of upper quadrant lymphedema secondary to cancer: clinical practice guideline from the Oncology Section of the American Physical Therapy Association. Phys Ther. 2017;97(7):729–745.
[5] Tamura MK. Introduction to patient-reported outcomes perspectives series. Clin J Am Soc Nephrol. 2017;12(11):1881–1881.
[6] Wu AW, Kharrazi H, Boulware LE, et al. Measure once, cut twice-adding patient-reported outcome measures to the electronic health record for comparative effectiveness research. J Clin Epidemiol. 2013;66(8 Suppl.):S12–S20.
[7] Prinsen CAC, Mokkink LB, Bouter LM, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147–1157.
[8] 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):737–745.
[9] Coriddi M, Dayan J, Sobti N, et al. Systematic review of patient-reported outcomes following surgical treatment of lymphedema. Cancers (Basel). 2020;12(3):565.
[10] 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):134–139.
[11] 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.
[12] Cemal Y, Jewell S, Albornoz CR, et al. Systematic review of quality of life and patient reported outcomes in patients with oncologic related lower extremity lymphedema. Lymphat Res Biol. 2013;11(1):14–19.
[13] Mokkink LB, Prinsen CA, Patrick DL, et al. COSMIN methodology for systematic reviews of patient-reported outcome measures (PROMs); user manual; 2018. Available from: https://www.cosmin.nl/tools/checklists-assessing-methodological-study-qualities/check-lists-assessing-methodological-study-qualities/
[14] 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):1171–1179.
[15] 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):1159–1170.
[16] Kleinlugtenbelt YV, Nienhuis RW, Bhandari M, et al. Are validated outcome measures used in distal radial fractures truly valid? A critical assessment using the COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) checklist. Bone Joint Res. 2016;5(4):153–161.
[17] Terwee CB, Mokkink LB, Knol DL, et al. Rating the methodological quality in systematic reviews of studies on measurement properties: a scoring system for the COSMIN checklist. Qual Life Res. 2012;21(4):651–657.
[18] Terwee CB, Bot SDM, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. J Clin Epidemiol. 2007;60(1):34–42.
[19] Prinsen CA, Vohra S, Rose MR, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" – a practical guideline. Trials. 2016;17(1):449.
[20] Augustin M, Conde Montero E, Hagenström K, et al. Validation of a short-form of the Freiburg Life Quality Assessment for lymphoedema (FLQA-LS) instrument. Br J Dermatol. 2018;179(6):1329–1333.
[21] Williams AE, Rapport F, Russell IT, et al. 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.
[22] Klernäs P, Johnsson A, Boyages J, et al. Test of responsiveness and sensitivity of the questionnaire "Lymphedema Quality of Life Inventory". Lymphat Res Biol. 2018;16(3):300–308.
[23] Klernas P, Johnsson A, Horstmann V, et al. Lymphedema Quality of Life Inventory (LyQLI)—development and investigation of validity and reliability. Qual Life Res. 2015;24(2):427–439.
[24] Klernas 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):135–145.
[25] Borman P, Yaman A, Denizli M, et al. 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 Rehab. 2018;64(3):205–212.
[26] Keeley V, Crooks S, Locke J, et al. A quality of life measure for limb lymphoedema (LYMQOL). J Lymphoedema. 2010;5(1):26–37.
[27] Wedin M, Fredrikson M, Ahlner E, et al. Validation of the Lymphoedema Quality of Life Questionnaire (LYMQOL) in Swedish cancer patients. Acta Oncol. 2020;59(3):365–371.
[28] Karayurt O, Deveci Z, Eyigor S, et al. Adaptation of quality of life measure for limb lymphedema-arm in Turkish women with breast cancer-related lymphedema. Cancer Nurs. 2019;28:2828.
[29] van de Pas CB, Biemans AA, Boonen RS, et al. Validation of the Lymphoedema Quality-of-Life Questionnaire (LYMQOL) in Dutch patients diagnosed with lymphoedema of the lower limbs. Phlebology. 2016;31(4):257–263.
[30] Bakar Y, Tuğral 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.
[31] Borman P, Yaman A, Denizli M, et al. The reliability and validity of lymphedema quality of life questionnaire-leg in Turkish patients with lower limb lymphedema. Lymphat Res Biol. 2020;18(1):42–48.
[32] Augustin M, Bross F, Foldi E, et al. Development, validation and clinical use of the FLQA-I, a disease-specific quality of life questionnaire for patients with lymphedema. Vasa. 2005;34(1):31–35.

[33] 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):131–138.
[34] Launois R, Megnigbeto AC, Pocquet K, et al. A specific quality of life scale in upper limb lymphedema: the ULL-27 questionnaire. Lymphology. 2002;35(1-760):181–187.
[35] Ridner SH, Doersam JK, Stolldorf DP, et al. Development and validation of the lymphedema symptom intensity and distress survey-lower limb. Lymphat Res Biol. 2018;16(6):538–546.
[36] Ridner SH, Dietrich MS. Development and validation of the lymphedema symptom and intensity survey-arm. Support Care Cancer. 2015;23(10):3103–3112.
[37] 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;20:10–16.
[38] Fu MR, Axelrod D, Guth AA, et al. mHealth self-care interventions: managing symptoms following breast cancer treatment. mHealth. 2016;2:28.
[39] Deng J, Ridner SH, Murphy BA, et al. Preliminary development of a lymphedema symptom assessment scale for patients with head and neck cancer. Support Care Cancer. 2012;20(8):1911–1918.
[40] De Vrieze T, Gebruers N, Nevelsteen I, et al. Responsiveness of the lymphedema functioning, disability, and health questionnaire for upper limb lymphedema in patients with breast cancer-related lymphedema. Lymphat Res Biol. 2020;18(4):365–373.
[41] De Vrieze T, Frippiat J, Deltombe T, et al. Cross-cultural validation of the French version of the lymphedema functioning, disability and health questionnaire for upper limb lymphedema (Lymph-ICF-UL). Disabil Rehabil. 2020:1–8. DOI:10.1080/09638288.2020.1716271
[42] 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. Lymphat Res Biol. 2019;17(3):347–355.
[43] 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 Theory Pract. 2019;35(4):327–340.
[44] 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):1728–1732.
[45] Devoogdt N, Van Kampen M, Geraerts I, et al. Lymphoedema functioning, disability and health questionnaire (Lymph-ICF): reliability and validity. Phys Ther. 2011;91(6):944–957.
[46] Wang C-M, Lee S-Y, Hsu K-F, et al. 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):1–11.
[47] Kostanoğlu A, Mbata GB, Gökmen GY, et al. The lymphedema functioning, disability, and health questionnaire for lower limb lymphedema: translation, reliability, and validation study of the Turkish version. Turk Gogus Kalp Dama. 2017;25(4):586–591.
[48] Ferreira KR, Carvalho RBM, de Andrade MFC, et al. Translation and cross-cultural adaptation of the lymphoedema functioning, disability and health questionnaire for lower limb lymphoedema into Portuguese language. Rev Bras Ginecol Obstet. 2016;38(2):88–96.
[49] 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):705–721.
[50] Abu Sharour L. Psychometric evaluation of the Arabic version of the lymphedema life impact scale in breast cancer patients. Breast J. 2020;26(3):563–565.
[51] Orhan C, Uzelpasaci 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):375–383.
[52] 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.
[53] Degirmenci B, Tuzun S, Of NS, et al. Reliability and validity of Turkish version of lymphedema life impact scale. Turk J Phys Med Rehab. 2019;65(2):147–153.
[54] Haghighat S, Montazeri A, Zayeri F, et al. 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).
[55] 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):128–138.
[56] Huang W, Zhang L, Yan J. Psychometric evaluation of the Chinese version of the revised illness perception questionnaire for breast cancer-related lymphedema. Eur J Cancer Care (Engl). 2019;28(1):e12900.
[57] Buchan J, Janda M, Box R, et al. Exercise barriers self-efficacy: development and validation of a subcale for individuals with cancer-related lymphedema. Health Qual Life Outcomes. 2015;13(1):37.
[58] 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.
[59] 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):100–107.
[60] Rotenberg S, Ruthralingam M, Hnatiw B, et al. Measurement properties of the multiple errands test: a systematic review. Arch Phys Med Rehabil. 2020;101(9):1628–1642.
[61] van Zijl F, Mokkink LB, Haagsma JA, et al. Evaluation of measurement properties of patient-reported outcome measures after rhinoplasty: a systematic review. JAMA Facial Plast Surg. 2019;21(2):152–162.
[62] Speyer R, Kim JH, Doma K, et al. Measurement properties of self-report questionnaires on health-related quality of life and functional health status in dysphonia: a systematic review using the COSMIN taxonomy. Qual Life Res. 2019;28(2):283–296.
[63] Crudgington H, Rogers M, Morris H, et al. Epilepsy-specific patient-reported outcome measures of children's health-related quality of life: a systematic review of measurement properties. Epilepsia. 2020;61(2):230–248.
[64] Moher D, Hopewell S, Schulz KF, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2012;10(1):28–55.
[65] Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMJ. 2010;340:c332.
[66] COSMIN. Research & publications COSMIN; 2020; [cited 2020 Apr 6]. Available from: https://www.cosmin.nl/finding-right-tool/conducting-study-measurement-properties/
[67] Terwee CB, Prinsen CA, 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):1159–1170.
[68] U.S. Department of Health and Human Services Food and Drug Administration (FDA) CfDEaRC, Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH). Guidance for industry patient-reported outcome measures: use in medical product development to support labeling claims. Rockville (MD); 2009; [cited 2019 Dec 23]. Available from: https://www.fda.gov/media/77832/download

[69] Mokkink LB, Prinsen CAC, Bouter LM, et al. The COnsensus-based Standards for the selection of health Measurement INstruments (COSMIN) and how to select an outcome measurement instrument. Braz J Phys Ther. 2016;20(2):105–113.
[70] International Test Commission. The ITC guidelines for translating and adapting tests (2nd ed.). IJT. 2018;18(2):101–134.
[71] Borsa JC, Damásio BF, Bandeira DR. Cross-cultural adaptation and validation of psychological instruments: some considerations. Paidéia (Ribeirão Preto). 2012;22(53):423–432.
[72] Boateng GO, Neilands TB, Frongillo EA, et al. Best practices for developing and validating scales for health, social, and behavioral research: a primer. Front Public Health. 2018;6:149.
[73] Sharawi N, Klima L, Shah R, et al. Evaluation of patient-reported outcome measures of functional recovery following caesarean section: a systematic review using the consensus-based standards for the selection of health measurement instruments (COSMIN) checklist. Anaesthesia. 2019;74(11):1439–1455.
[74] Tsuchiya M. Development of the Arm Lymphoedema Physical Discomfort Scale: validity and reliability testing with Japanese breast cancer survivors. Japan J Health Psychol. 2012;25(1):74–82.
[75] Moher D, Liberati A, Tetzlaff J, et al. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.