

Responsiveness of the Lymphedema Functioning, Disability, and Health Questionnaire for Upper Limb Lymphedema in Patients with Breast Cancer-Related Lymphedema

Tessa De Vrieze, PT, MT,^{1,2} Nick Gebruers, PT,^{2,3} Ines Nevelsteen, MD,⁴ Wiebren A.A. Tjalma, MD,^{3,5,6} Sarah Thomis, MD,^{7,8} An De Groef, PT, MT,¹ Lore Dams, PT, MT,^{1,2} and Nele Devoogdt, PT^{1,7,8}

Abstract

Background: The Lymphedema Functioning, Disability, and Health questionnaire for Upper Limb Lymphedema (Lymph-ICF-UL) is a health-related quality-of-life questionnaire for patients with breast cancer-related lymphedema. Previous testing of this questionnaire showed very good psychometric properties; however, responsiveness has not yet been established. The aim of this study was to determine its internal and external responsiveness.

Methods and Results: Ninety-five patients treated with decongestive lymphatic therapy in a longitudinal trial were recruited. Patients completed the Lymph-ICF-UL twice within a time interval of 7 weeks (“intensive group” receiving intensive treatment; $n=73$) or 3 months (“stable group” receiving maintenance treatment; $n=22$), and once the Global Perceived Effect (GPE) questionnaire at the second time point. The significance of change in scores and standardized response mean (SRM) were determined for the total and domain scores. Correlations between Lymph-ICF-UL and GPE were ascertained. In addition, the minimal clinical important difference (MCID) was determined. The Lymph-ICF-UL total score changed significantly in the intensive group ($p<0.001$) and nonsignificantly for those in the stable group ($p=0.25$). The SRM represented moderate responsiveness (0.65). Patients who reported a clinical improvement (=responders) after intensive treatment showed a significant decrease in total score ($p<0.001$), this was also the case for nonresponders ($p<0.001$). Lymph-ICF-UL total and domain scores showed nonsignificant weak correlations with the GPE ($p>0.05$). There was a significant difference in mean total score changes between responders and nonresponders ($p<0.001$). MCID for the total score was 9%.

Conclusion: The Lymph-ICF-UL is responsive to change after decongestive lymphatic therapy. No correlations were found between Lymph-ICF-UL change scores and GPE. Future studies should be conducted in a clinical setting, with more variability between participants and their treatment responses.

Keywords: breast neoplasms, lymphedema, Lymph-ICF-UL questionnaire, responsiveness, psychometric properties

Introduction

LYMPHEDEMA IS A troublesome morbidity affecting ~17% of the women treated for breast cancer.¹ The edema can be measured objectively with different valid and reliable assessment methods (water displacement, circum-

ference measurements, etc.).² However, an objective assessment of the volume lacks the power to encounter the real burden of lymphedema.

Besides swelling, patients can suffer from problems in physical, social, and mental functioning.^{3,4} In addition, breast cancer-related lymphedema (BCRL) lowers the quality of

¹Department of Rehabilitation Sciences, KU Leuven—University of Leuven, Leuven, Belgium.

²Department of Rehabilitation Sciences and Physiotherapy, MOVANT, University of Antwerp, Antwerp, Belgium.

³Multidisciplinary Oedema Clinic, University of Antwerp and Antwerp University Hospital, Antwerp, Belgium.

⁴Multidisciplinary Breast Centre, UZ Leuven—University Hospitals Leuven, Leuven, Belgium.

⁵Department of Molecular Imaging, Pathology, Radiotherapy, Oncology (MIPRO), Faculty of Medicine and Health Sciences, University of Antwerp, Antwerp, Belgium.

⁶Multidisciplinary Breast Clinic, Antwerp University Hospital, Antwerp, Belgium.

Departments of ⁷Vascular Surgery and ⁸Physical Medicine and Rehabilitation, Centre for Lymphedema, UZ Leuven—University Hospitals Leuven, Leuven, Belgium.

life.^{5,6} Therefore, the Lymphedema Functioning, Disability, and Health questionnaire for the upper limb (Lymph-ICF-UL) was developed to assess these impairments.⁷ This questionnaire aims to quantify impairments in function, activity limitations, and participation restrictions that are related to upper limb lymphedema. In contrast to other lymphedema-related questionnaires, it is based on the terminology of the International Classification of Functioning, Disability and Health (ICF) of the World Health Organization (WHO).⁸

The quality and usefulness of a questionnaire are determined by its psychometric properties, including validity, reliability, and responsiveness. The reliability and validity of the Lymph-ICF-UL have already been examined in patients with BCRL and showed very good (reliability) to good (validity) psychometric parameters.^{7,9} However, responsiveness of the Lymph-ICF-UL has yet to be ascertained.

Therefore, the aim of this study was to examine the internal and external responsiveness of the Lymph-ICF-UL in patients with BCRL.

Materials and Methods

Study design

Participants of the EForT-BCRL trial¹⁰ were recruited in the University Hospitals of Leuven, Antwerp University Hospital, Ghent University Hospital and General Hospital Groeninge in Kortrijk, Belgium. Approval for this study was obtained by the Ethical Committee of the University Hospitals of Leuven (main Ethical Committee), as well as of the ethical committees of all other participating centers (CME reference S58689, EudraCT Number 2015-004822-33). This study was reported following the COSMIN (CONsensus-based Standards for the selection of health Measurement INSTRuments) guidelines.¹¹

Participants

Female patients with BCRL of the arm and/or hand, who were about to start the intensive phase of decongestive lymphatic therapy through participation in the EForT-BCRL trial (= intensive group), or participants who were at least 3 months in the trial's maintenance phase of this study (= stable group), were recruited. Criteria for inclusion and exclusion are presented in Table 1. Note that no formal power analysis has been performed, but that the sample size was completely determined by the response rate of the ongoing EForT-BCRL trial.¹⁰

Study procedure

This study was conducted between March 2016 and October 2018. All patients provided written informed consent before treatment. Descriptive data (participant's age, body mass index, type of breast surgery, side of surgery, hand dominance, type of adjuvant treatment (radiotherapy, chemotherapy, hormone therapy, or immunotherapy), duration, and stage of lymphedema) were collected by interviewing the participants and by consulting their medical record.

Figure 1 shows the study procedure. Patients were asked to complete the Lymph-ICF-UL questionnaire before the start of the intensive treatment phase (intensive group) or at the beginning of month 3 of their maintenance phase (stable group). In addition, the volume of each of the participants' arms was determined by circumference measurements using a perimeter after which the total arm volume was calculated using the truncated cone formula.^{12,13}

The intensive treatment lasted for 3 weeks and consisted of all components of the decongestive lymphatic therapy, as recommended in the consensus document of the International Society of Lymphedema (ISL): manual lymph drainage, skin care, exercises, and multilayer bandaging. When the volume of the arm was decreased optimally and pitting was absent, a compression stocking and glove were measured. When patients received the compression garment, the 6 months lasting maintenance phase started, consisting of: manual lymph drainage, skin care, exercises, and wearing the compression garment.^{10,14}

The second time point for data collection was performed 4 weeks after wearing the stocking (intensive group) in order patients could get used to the feeling of wearing a stocking and/or gauntlet, or at the end of month 6 of their maintenance phase (stable group). Again, this second evaluation consisted of completing the Lymph-ICF-UL questionnaire, this time together with the Global Perceived Effect (GPE) questionnaire (Fig. 1).

All treatments and assessments were provided at the Department of Physical Medicine and Rehabilitation of the University Hospitals of Leuven, at the Multidisciplinary Breast Clinic of the Antwerp University Hospital, at the Departments of Plastic Surgery and Radiotherapy of the Ghent University Hospital, and at the Centre for Oncology at the General Hospital Groeninge in Kortrijk. Measurements were performed by one of three physical therapists, specialized in edema therapy (S.V.D.B., L.V., and T.D.V.) who were blinded for the treatment allocation of patients. Treatments

TABLE 1. INCLUSION AND EXCLUSION CRITERIA

<i>Inclusion criteria</i>	<i>Exclusion criteria</i>
Female	Known local recurrence or metastasis
Diagnosis of breast cancer	Cognitive limitations
Surgery, chemotherapy, radiation therapy completed for at least 3 months	No BCRL/other type of edema
Unilateral arm and/or hand lymphedema (>5% difference at the time of inclusion in the EForT-BCRL trial, adjusted for limb dominance)	
Starting an intensive decongestive treatment including multilayer bandaging, as part of the EForT-BCRL trial (intensive group) or being in the maintenance phase of the EForT-BCRL trial for at least 3 months (stable group)	
Native Dutch speaking or being able to read, write, and understand the Dutch language	

BCRL, breast cancer-related lymphedema.

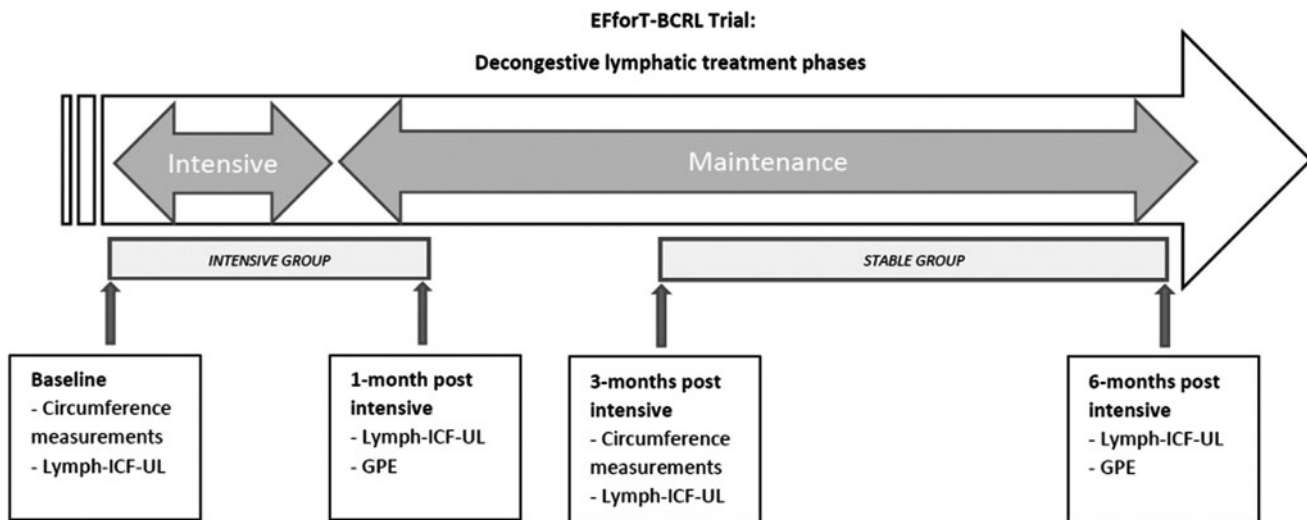


FIG. 1. Illustration of the study procedure.

were performed by one of four physical therapists, specialized in edema therapy (L.B., R.V.H., L.V., T.D.V.).

Outcome variables

Following questionnaires and measurements were used.

Lymphedema Functioning, Disability, and Health Questionnaire for upper limb lymphedema^{7,9}. The Lymph-ICF-UL is a self-reported comprehensive evaluation tool assessing impairments in functioning, activity limitations, and participation restrictions in patients with BCRL during a 2-week recall period. It consists of 29 questions, covering five domains: physical function, mental function, household activities, mobility activities, and life and social activities. Each of the questions has to be scored on an 11-point Likert scale with a score between 0 and 10. The total score on the Lymph-ICF-UL and the scores on the five domains range between 0 and 100. The higher the score, the more problems patients experience. Reliability and content validity have shown to be very good; construct validity is good.^{7,9}

GPE questionnaire. The reference criterion used in this study to investigate external responsiveness was the GPE scale.¹⁵ The GPE is a patient-reported outcome measure stating the amount of improvement as perceived by the patient. The following question was asked to the patients of the intensive group: “To what extent did you recover from your lymphedema-related symptoms and complaints since the beginning of the treatment?” Alternatively, to the patients of the stable group, the following question was asked: “To what extent have your lymphedema-related symptoms and complaints (what you feel, can perform, etc) changed compared to the previous evaluation moment within the EfforT-BCRL study? This means: we want to know the degree of change in your complaints, only between(date previous assessment) and(today) (not in comparison with the period before your participation in the study).”

It measures the perception of the patient with use of an ordinal scale. A 7-point Likert scale ranging from (1) very much better, (2) much better, (3) a little better, (4) unchanged,

(5) a little worse, (6) much worse to (7) very much worse was used, as recommended by Ostelo.¹⁶ Literature shows that scores 1 and 2 can be considered as a clinically relevant improvement,¹⁷ whereas a score of 3 (a little better) should be considered as unchanged as this reflects a minimum degree of improvement that could be experienced in patients just by being treated with attention for the current health-related problems.^{16,18} Consequently, patients scoring the GPE with 1 or 2 (very much better to much better) are further referred to as “responders,” whereas patients scoring the GPE with scores 3, 4, 5, 6, or 7 (a little better to very much worse) are further referred to as “nonresponders.”^{16,19–22} The GPE proved to have an excellent reproducibility.^[23]

Investigation of responsiveness

There is no consensus on how responsiveness of measures should be quantified and it is further complicated by the multiple definitions that are used.²⁴ In general, literature suggests that there are two major facets of responsiveness: internal and external responsiveness. Internal responsiveness characterizes the ability of a measure to show changes within a particular period of time.²⁴ The observed changes in the measures are attributed to clinically relevant changes in health.²⁴ Consequently, the internal responsiveness of a measure will depend upon the particular treatment that is provided to patients as well as the specific outcomes that are described to determine treatment efficacy.²⁴

In addition, external responsiveness reflects the extent to which changes in a measure over a particular period of time relate to corresponding changes in an external reference measure of a person’s health status.²⁴ In this type of responsiveness, the measure itself is not of primary interest, but the relationship between change in the measure and change in the external standard.²⁴ In contrast to internal responsiveness, the external responsiveness of a measure will solely depend on the choice of the external reference measure and not on the investigated treatment.²⁴

To investigate the internal and external responsiveness of the Lymph-ICF-UL, we were interested in the following topics for which we had formulated subsequent hypotheses:

Internal responsiveness.

1. Whether or not the Lymph-ICF-UL could demonstrate a statistically significant change in scores before and after the intensive treatment phase.

Hypothesis 1: In the intensive group, there would be a statistically significant change in mean total scores of the Lymph-ICF-UL between the two evaluation moments ($p < 0.05$).

2. Whether or not the Lymph-ICF-UL could demonstrate a statistically significant change in scores before and after the 3 months of maintenance treatments.

Hypothesis 2: In the stable group, there would be no statistically significant difference in mean total scores of the Lymph-ICF-UL between the two evaluation moments ($p > 0.05$).

3. Whether or not the Lymph-ICF-UL is able to show a relatively small level of variability in change scores in relation to the average change in scores between the two evaluation moments, by means of the standardized response mean (SRM) as an effect size.

Hypothesis 3: The calculated SRM values reflecting the variability of the change scores of the Lymph-ICF-UL would represent moderate (≥ 0.50) to large (≥ 0.80) responsiveness for the Lymph-ICF-UL total score.²⁴

External responsiveness.

4. Whether the Lymph-ICF-UL could demonstrate a statistically significant change in scores before and after intensive therapy in the responders on the one hand and in the nonresponders on the other hand. In addition whether the Lymph-ICF-UL could demonstrate a statistically significant difference in mean change score between responders and nonresponders after intensive treatments.

Hypothesis 4: The change in mean Lymph-ICF-UL total score before and after intensive treatment would be statistically significantly different between responders and nonresponders ($p < 0.05$).

5. Whether the Lymph-ICF-UL would show a correlation between the change in scores (before and after intensive/maintenance treatments) and the GPE.

Hypothesis 5: There would be at least a moderate correlation (≥ 0.3) between the change in mean Lymph-ICF-UL scores (of both the intensive and the stable group together) and the score on the GPE.

6. The minimal clinical important difference (MCID) of the Lymph-ICF-UL.

Hypothesis 6: The MCID for responders on the total score of the Lymph-ICF-UL would be < 10 (10%).

Statistical analyses and interpretation

Data analyses were performed using the Statistical Package for Social Sciences 25 for Windows (SPSS, Inc., Chicago, IL). Normality of the variables was tested using the One-Sample Kolmogorov–Smirnov test and descriptive statistics were calculated. The 0.05 level of significance was applied.

Data are presented as number and percentage for categorical variables and mean with standard deviation (SD) (normal distribution) or median with interquartile range (non-normal distribution) for continuous variables, unless otherwise stated.

Considering the investigation of internal responsiveness, the following statistical tests were performed:

1. Wilcoxon-signed-rank tests were used to determine whether the Lymph-ICF-UL total and domain scores were significantly different before and after the intensive treatment phase.
2. Also to determine whether the Lymph-ICF-UL total and domain scores were significantly different between the two evaluation moments during the maintenance phase, the Wilcoxon-signed-rank tests were performed.
3. The SRM as effect size was calculated for the responsive group using the following formula²⁴:
$$\frac{\text{difference in Lymph-ICF-UL mean scores}}{\text{standard deviation of the difference in mean scores}}$$
 SRM values of 0.20, 0.50, and 0.80 or higher have been proposed to represent small, moderate, and large responsiveness, respectively.^{24–27}

Considering the investigation of external responsiveness, the following statistical tests were performed:

4. Wilcoxon-signed-rank tests were used to determine whether the Lymph-ICF-UL total and domain scores were significantly different before and after the intensive phase of treatment in the responders and nonresponders group separately. Comparability of the group responders and nonresponders was tested with Mann–Whitney-*U* for numeric data and with chi-square for categorical data. In addition, to investigate significant differences in the mean change in scores between responders and nonresponders before and

TABLE 2. CHARACTERISTICS OF THE PARTICIPANTS ($N = 95$)

Variable	Outcome
Age (years)	62 (10)
Body mass index (kg/m ²)	28.7 (5.6)
Lymphedema volume arm (absolute difference) (mL)	540 (388)
Hand circumference (absolute difference) (cm)	3.0 (12.1)
Duration of lymphedema (months)	53.0 (42.5)
BCRL stages, n (%)	
I	12 (12.6)
IIa	56 (59.0)
IIb	27 (28.4)
Breast surgery, n (%)	
Mastectomy	44 (46.3)
Breast-conserving surgery	51 (53.7)
Axillary clearance, n (%)	
Sentinel lymph node biopsy alone	3 (3.2)
Axillary lymph node dissection	92 (96.8)
Surgery at the dominant side, n (%)	41 (43.2)
Radiotherapy, n (%)	92 (96.8)
Chemotherapy, n (%)	81 (85.3)
Hormone therapy, n (%)	77 (81.1)
Immunotherapy (Herceptin), n (%)	21 (22.1)

BCRL stages as described by the International Society of Lymphology; descriptive statistics are depicted as mean (SD). SD, standard deviation.

TABLE 3. MEAN PRE- AND POSTTREATMENT SCORES OF THE LYMPH-ICF-UL IN BOTH THE INTENSIVE AND STABLE GROUP

Score	Intensive group (n=73)			Stable group (n=22)		
	Mean score pre	Mean score post	p	Mean score pre	Mean score post	p
Lymph-ICF-UL total score	38	28	<0.001**	27	31	0.25
Physical function score	44	25	<0.001**	26	30	0.32
Mental function score	31	20	<0.001**	19	24	0.33
Household activities score	41	33	<0.001**	30	30	1.00
Mobility activities score	37	32	0.06	32	39	0.08
Life and social activities score	35	29	0.03*	26	29	0.43

* p -value <0.05, ** p -value <0.01.

Lymph-ICF-UL, Lymphedema Functioning, Disability, and Health questionnaire for the upper limb.

after intensive therapy, a two-way analysis of variance (ANOVA) for repeated measures statistic was applied.

- Spearman's rho correlation analysis was performed on the entire group (both the intensive and stable groups) to determine the correlation between the change in Lymph-ICF-UL scores and the reported GPE. According to Cohen, the correlations required values of 0.3 or higher to be regarded as a good anchor.^{28,29}
- The MCID represents the smallest change in score that the participant perceives as a meaningful improvement.³⁰ If a participant's score is above the MCID, it is considered clinically relevant. To define the MCID, the mean change scores on the Lymph-ICF-UL of the participants who reported an important clinically important improvement (responders scoring the GPE with 2, i.e., "much better") were used.³⁰ Consequently, to investigate the MCID, descriptive statistics were used to describe the mean (\pm SD) of the total and domain scores corresponding to the responders of the entire group (both the intensive and stable groups).

Results

In this study, 95 participants were recruited. Of these, 73 participants were enrolled in the intensive group and 22 participants in the stable group. The mean age of the participants was 62 (\pm 10) years and mean body mass index was 29 (\pm 6) kg/m². In the intensive group ($n=73$), the mean absolute difference in lymphedema volume of the arm was 541 (\pm 481) mL. In the stable group ($n=22$), the mean absolute difference in lymphedema volume of the arm was 384 (\pm 282) mL. The characteristics of the participants are presented in Table 2.

Internal responsiveness

Change in Lymph-ICF-UL scores after treatment: intensive group. The mean pre- and postintensive treatment scores of the Lymph-ICF-UL are given in Table 3.

A statistically significant difference ($p < 0.05$) was present in the intensive group between the pre- and postintensive treatment total scores of the Lymph-ICF-UL as well as in all domain scores, except for the mobility activities domain ($p = 0.06$).

Change in Lymph-ICF-UL scores after treatment: stable group. The mean pre- and postmaintenance treatment scores of the Lymph-ICF-UL are presented in Table 3. There was no statistically significant difference in total nor domain scores between the two assessments ($p > 0.05$).

Effect size: SRM. SRM values are presented in Table 4. The effect size of the total score represented moderate responsiveness (0.65). Highest SRM value is shown in the physical functions domain (0.99), representing good responsiveness. Lowest value was for the mobility activities domain, showing small responsiveness (0.21).

External responsiveness

Difference in pre- and postintensive treatment scores for responders and nonresponders. Table 5 presents an overview of the mean total and domain scores of the Lymph-ICF-UL before and after intensive treatment, as well as the mean change scores before and after intensive treatment for/between responders and nonresponders. Responders ($n=39$) showed a statistically significant decrease in the Lymph-ICF-UL total score, physical function, mental function, and mobility activities domain scores over time ($p < 0.05$). Other domains were not significantly different before and after intensive treatment. Nonresponders ($n=34$) showed a statistically significant decrease in the Lymph-ICF-UL total score, physical function, mental function, and household activities domain scores ($p < 0.05$).

Preintensive treatment scores on the Lymph-ICF-UL were significantly different in both groups for the total score ($p = 0.02$) as for the domains physical function ($p = 0.01$), household activities ($p = 0.08$), and life and social activities

TABLE 4. STANDARDIZED RESPONSE MEANS CALCULATED FOR THE INTENSIVE GROUP ($n=73$)

Lymph-ICF-UL	Total	Physical functions	Mental functions	Household activities	Mobility activities	Life and social activities
SRM	0.65	0.99	0.54	0.36	0.21	0.27

SRM, standardized response mean.

TABLE 5. SEPARATE MEAN PRE- AND POSTTREATMENT SCORES OF LYMPH-ICF-UL FOR RESPONDERS AND NONRESPONDERS AND THE MEAN CHANGE IN SCORES BEFORE AND AFTER INTENSIVE TREATMENT BETWEEN RESPONDERS AND NONRESPONDERS

Intensive group	Responders (n=39)				Nonresponders (n=34)				
	Mean score pre (SD)	Mean score post (SD)	Difference (SD)	p-value (within group)	Mean score pre (SD)	Mean score post (SD)	Difference (SD)	p-value (within group)	p-value mean change between groups (ANOVA)
Lymph-ICF-UL total score	32 (21)	21 (18)	11 (16)	<0.001**	44 (21)	36 (20)	8 (14)	<0.001**	<0.001**
Physical function score	36 (20)	19 (17)	17 (15)	<0.001**	52 (25)	32 (22)	20 (22)	<0.001**	<0.001**
Mental function score	26 (26)	15 (21)	11 (19)	0.001**	37 (32)	26 (25)	11 (22)	0.006**	<0.001**
Household activities score	32 (27)	25 (23)	7 (22)	0.092	51 (28)	41 (28)	10 (25)	0.033*	0.010**
Mobility activities score	32 (25)	24 (22)	8 (26)	0.028*	42 (24)	41 (24)	1 (19)	0.662	0.093
Life and social activities score	30 (27)	23 (21)	7 (25)	0.142	41 (26)	36 (26)	5 (20)	0.155	0.030*

*p-value <0.05, **p-value <0.01. ANOVA, analysis of variance.

($p=0.04$) domain scores of the Lymph-ICF-UL, in which the nonresponders showed relatively more problems in functioning at baseline compared with the responders.

The mean change in scores before and after intensive treatment was significantly different between responders and nonresponders for the total score ($p<0.001$), physical function ($p<0.001$), mental function ($p<0.001$), household activities ($p=0.01$), and life and social activities ($p=0.03$) scores (Table 5).

Correlations between change scores and GPE. Correlations between changes in scores of the Lymph-ICF-UL (Δ -Lymph-ICF-UL) and the GPE scores are given in Table 6. The scores of all the 95 participants were used. The Lymph-ICF-UL total score as well as the physical functions, mental functions, household activities, and mobility activities domains showed nonsignificant weak positive correlations with the GPE.

MCID in Lymph-ICF-UL score. An overview of the MCIDs (SDs) associated with the Lymph-ICF-UL total and domain scores is provided in Table 7. The MCID estimate for the Lymph-ICF-UL total score was 9%, physical function 14%, mental function 7%, household activities 8%, mobility activities 6%, and life and social activities 5%.

Discussion

The aim of this study was to investigate the responsiveness of the Lymph-ICF-UL questionnaire.

As an answer to the methodological inconsistencies in the literature regarding responsiveness, the COSMIN panel reached consensus on the definitions of measurement properties for health-related patient-reported outcomes in an international Delphi study.³¹ A checklist was developed to determine the methodological quality of studies on measurement properties.^{32,33} The definition of responsiveness according to the COSMIN initiative is “the ability of a health-related patient reported outcome instrument to detect change over time in the construct to be measured.”³¹ In this study, this was translated to the ability of the Lymph-ICF-UL to detect a clinically important change in number of problems in functioning of patients with BCRL, as part of the external responsiveness of the questionnaire.

Results showed that, in this study, only one out of the six hypotheses regarding the internal and external responsiveness was rejected. All three hypotheses regarding internal responsiveness (hypotheses 1–3) were accepted. There was a

TABLE 6. SPEARMAN'S RHO CORRELATIONS BETWEEN CHANGE SCORES OF LYMPH-ICF AND GLOBAL PERCEIVED EFFECT SCORES (N=95)

	Domain	r_s	p
Δ Lymph-ICF-UL	Total	0.134	0.20
	Physical functions	0.092	0.37
	Mental functions	0.164	0.11
	Household activities	0.112	0.28
	Mobility activities	0.195	0.06
	Life and social activities	-0.041	0.70

Δ Lymph-ICF-UL, mean change of scores of Lymph Functioning, Disability, and Health questionnaire for upper limb lymphedema.

TABLE 7. OVERVIEW OF THE MINIMAL CLINICAL IMPORTANT DIFFERENCES (STANDARD DEVIATIONS) OF THE LYMPH-ICF-UL SCORES ACCORDING TO THE DIFFERENT SCORES ON THE GLOBAL PERCEIVED EFFECT (N=95)

GPE	<i>Lymph-ICF-UL domains</i>					
	<i>Total score</i>	<i>Physical functions</i>	<i>Mental function</i>	<i>Household activities</i>	<i>Mobility activities</i>	<i>Life and social activities</i>
1 = very much better	n = 8 7 (6)	n = 8 7 (13)	n = 8 17 (18)	n = 8 10 (16)	n = 8 4 (11)	n = 8 2 (9)
2 = much better (MCIDs)	n = 37 9 (18)	n = 37 14 (17)	n = 37 7 (21)	n = 37 8 (23)	n = 37 6 (28)	n = 36 5 (28)
3 = a little better	n = 31 9 (12)	n = 31 20 (22)	n = 31 14 (18)	n = 31 6 (19)	n = 31 2 (15)	n = 31 3 (18)
4 = the same	n = 15 3 (17)	n = 15 9 (24)	n = 15 3 (18)	n = 15 6 (32)	n = 15 3 (23)	n = 15 7 (20)
5 = worse	n = 4 17 (19)	n = 4 20 (24)	n = 4 18 (27)	n = 4 8 (14)	n = 4 19 (23)	n = 4 12 (20)

Scores are depicted as mean (SD).
MCID, minimal clinical important difference.

statistically significant difference in the intensive group between the pre- and postintensive treatment total scores of the Lymph-ICF-UL as well as in all domain scores, except for the mobility activities domain (Hypothesis 1). A reason for this might be that restrictions in mobility activities can be influenced by other factors as well, besides BCRL. Furthermore, there was no statistically significant difference in total nor domain scores between the two evaluation moments in the maintenance phase (Hypothesis 2). Finally, the effect size of the total score represented moderate responsiveness (SRM 0.65), suggesting that the questionnaire is able to indicate a clinically meaningful change in total score²⁴ (Hypothesis 3).

Regarding external responsiveness, 2 out of 3 hypotheses (hypotheses 4 and 6) were accepted. We expected that there would be a statistically significant difference between the change in total score of the responders and nonresponders, which was confirmed by this study (Hypothesis 4). Only for the mobility activities domain, this change score was non-significantly different.

Responders showed a statistically significant decrease in total score as well as in most of the domain scores of the Lymph-ICF-UL after the intensive treatments. However, surprisingly this was also the case for nonresponders. Whereas our analyses revealed that the nonresponders showed a relatively higher level of problems in functioning at baseline than the responders. Consequently, in patients with a higher number of problems in functioning at baseline, a relatively greater improvement in functioning after treatment can be expected.

Furthermore, when correlating the change scores of the Lymph-ICF-UL with the reported GPE, we expected at least moderate correlations (Hypothesis 5). However, results revealed mainly nonsignificant weak positive correlations. Therefore, this hypothesis could not be accepted. A major drawback for this correlation analysis was the relatively under-representation of patients who reported to be deteriorated after their treatment. Whereas, while designing this study, we decided to include also a subgroup of patients who were currently at the end of the maintenance phase (the stable group) as an attempt to encounter this and to include also patients who might have been deteriorated. Nevertheless, results indicated that, of the entire group, only 4% reported to

be worse, 16% reported to be unchanged, 33% reported to have a little bit improved, and 47% reported to have improved (much or very much better). Consequently, the majority of participants reported to be improved or unchanged, resulting in a rather homogeneous study sample. Other authors have discussed the use of GPE as an anchor as it might be very dependent on the status of a patient and, therefore, it might be more a measure of the patient's status than of the change in health status over time.^{23,34}

Lastly, we hypothesized that the MCID for responders on the total score of the Lymph-ICF-UL would be <10 (10%) (Hypothesis 6), which was an arbitrary chosen cutoff point based on empirical experience. As the total score of the Lymph-ICF-UL represented an MCID of 9%, our last hypothesis could be accepted as well. This result entails that if this total score decreases with at least 9 on 100, an overall clinical improvement will be experienced. This MCID exceeds the earlier reported standard error of measurement of 5, thereby eliminating the possibility that the change in score could be due to any measurement error.⁹ Consequently, the reported limitation that the MCID does not take measurement precision into account³⁵ is partially compensated this way.

Limitations and strengths

A strength of this study is that, in the investigation on responsiveness, the recommendations of the COSMIN panel were taken into account. An integrated system making use of multiple methods to define internal and external anchor-based responsiveness was applied.

Some limitations should be considered. First, as this investigation was conducted on a cohort of participants of the EforT-BCRL trial, patient characteristics, protocol, and treatment outcomes were rather homogeneous as the majority of the participants indicated to have improved or not to have changed, which might have induced a selection bias. There was a lack of participants who reported a deterioration (only 4 out of 95 participants), which was a shortcoming for the purpose of this investigation.

Second, the moment of completion of the questionnaires was for each patient at the end of a 1-hour clinical assessment.

This could have entailed an influence on patient's motivation and concentration level to spend some extra time and effort on reading every question with full attention. For this reason, we might suspect that some interpretation errors of the scoring system could have occurred (for instance when the anchors "very well" and "not at all" were converted in some questions but was not noticed by the participant). However, this was explained in advance. Last, as completion of the questionnaires occurred at the end of fixed evaluation moments in accordance with the EForT-trials' protocol, the time in-between the two evaluations was different for the intensive group (7 weeks) as for the stable group (3 months). However, we believe this has not affected our study results.

Implications for clinical practice

The Lymph-ICF-UL questionnaire already proved to be appropriate and useful in clinical practice by showing very good reliability (intra-class correlation coefficients between 0.79 and 0.95 and Cronbach alpha coefficients higher than 0.80), very good face and content validity, and good construct validity (79% of accepted hypotheses regarding convergent/divergent validity).⁹ This study reveals that the Lymph-ICF-UL is sensitive to detect changes over time. A change of 9% in total score indicates a clinically relevant change in the number of problems in functioning, of a patient with BCRL. A change of 14% in the physical function domain score designates a clinically relevant change in the number of problems regarding physical functions. Furthermore, a change of 7% in the mental function domain score indicates a clinically relevant change in the number of problems regarding mental functions. In the household activities domain, a change of 8% describes a clinically relevant change in the number of problems regarding household activities. Likewise, in the mobility activities domain, a change of 6% indicates a clinically relevant change in the number of problems regarding mobility activities. Lastly, in the life and social activities domain, a change of 5% indicates a clinically relevant change in the number of participation problems.

Recommendations for future research

When correlating the Lymph-ICF-UL scores with the GPE, the questionnaire showed a reduced ability to discriminate between the number of changes in Lymph-ICF-UL scores and the actual clinical improvement as reported by participants. As in our opinion, this is mainly due to the strict protocol in which this investigation occurred, a future investigation should be continued in a clinical setting, resulting in more variability between the study participants and consequently in their treatment responses. Furthermore, attention should be paid on the moment of completion of the questionnaires in order patients to be fully concentrated.

Conclusion

This study revealed that the Lymph-ICF-UL questionnaire is responsive to change after decongestive lymphatic therapy, in patients with BCRL. Based on the GPE as anchor-based method, an MCID of 9% indicates a clinically relevant change. No correlation between Lymph-ICF-UL change scores and GPE was found. Future studies should be conducted in a clinical

setting, enabling a greater amount of variability between the study participants and treatment responses.

Acknowledgments

The authors are very grateful to the hospitals collaborating in this study. The authors also extend grateful thanks to the study participants, to Lien Billiet, Roxane Van Hemelrijck, and Lore Vos for performing the treatments and to Shanah Van den Bosch and Lore Vos for measuring and handing over the questionnaires to the study participants. All authors critically revised the article for important intellectual content and approved the final article. A.D.G. is a postdoctoral research fellow of the FWO-Flanders.

Clinical Trial Registration Number

The study is part of a double-blind multicenter randomized controlled trial (EForT-BCRL trial), which is registered in clinicaltrials.gov (NCT02609724).

Author Disclosure Statement

No competing financial interests exist.

Funding Information

This study is financed by the Agency for Innovation by Science and Technology, Applied Biomedical Research (IWT 150178). To arrange such financing, a separate collaboration agreement has been signed by the KU Leuven and the beneficiaries.

References

- DiSipio T, Rye S, Newman B, Hayes S. Incidence of unilateral arm lymphoedema after breast cancer: A systematic review and meta-analysis. *Lancet Oncol* 2013; 14:500–515.
- Hidding JT, Viehoff PB, Beurskens CH, van Laarhoven HW, Nijhuis-van der Sanden MW, van der Wees PJ. Measurement properties of instruments for measuring of lymphedema: Systematic review. *Phys Ther* 2016; 96: 1965–1981.
- Fu MR, Ridner SH, Hu SH, Stewart BR, Cormier JN, Armer JM. Psychosocial impact of lymphedema: A systematic review of literature from 2004 to 2011. *Psycho-oncology* 2013; 22:1466–1484.
- Bojinovic-Rodic D, Popovic-Petrovic S, Tomic S, Markez S, Zivanic D. Upper extremity function and quality of life in patients with breast cancer related lymphedema. *Vojnosanit Pregl* 2016; 73:825–830.
- Ridner SH, Dietrich MS. Development and validation of the Lymphedema Symptom and Intensity Survey-Arm. *Support Care Cancer* 2015; 23:3103–3112.
- Lopez Penha TR, van Bodegraven J, Winkens B, Heuts EM, Voogd AC, von Meyenfeldt MF. The quality of life in long-term breast cancer survivors with breast cancer related lymphedema. *Acta Chir Belg* 2014; 114:239–244.
- Devoogdt N, Van Kampen M, Geraerts I, Coremans T, Christiaens MR. Lymphoedema Functioning, Disability and Health questionnaire (Lymph-ICF): Reliability and validity. *Phys Ther* 2011:944–957.
- International Classification of Functioning, Disability and Health: ICF*. Geneva, Switzerland: World Health Organization; 2001.

9. 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; 17:347–355.
10. De Vrieze T, Vos L, Gebruers N, et al. Protocol of a randomised controlled trial regarding the effectiveness of fluoroscopy-guided manual lymph drainage for the treatment of breast cancer-related lymphoedema (EFforT-BCRL trial). *Eur J Obstet Gynecol Reprod Biol* 2017; 221: 177–188.
11. Mokkink LB, Terwee CB, Knol DL, et al. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Med Res Methodol* 2010; 10:22.
12. Taylor R, Jayasinghe UW, Koelmeyer L, Ung O, Boyages J. Reliability and validity of arm volume measurements for assessment of lymphedema. *Phys Ther* 2006; 86:205–214.
13. Devoogdt N, Lemkens H, Geraerts I, et al. A new device to measure upper limb circumferences: Validity and reliability. *Int Angiol* 2010; 29:401–407.
14. The Diagnosis and Treatment of Peripheral Lymphedema: 2016 consensus document of the International Society of Lymphology. *Lymphology* 2016; 49:170–184.
15. Hudak PL, Wright JG. The characteristics of patient satisfaction measures. *Spine* 2000; 25:3167–3177.
16. Ostelo RW, de Vet HC. Clinically important outcomes in low back pain. *Best Pract Res Clin Rheumatol* 2005; 19: 593–607.
17. Farrar JT, Portenoy RK, Berlin JA, Kinman JL, Strom BL. Defining the clinically important difference in pain outcome measures. *Pain* 2000; 88:287–294.
18. Salaffi F, Stancati A, Silvestri CA, Ciapetti A, Grassi W. Minimal clinically important changes in chronic musculoskeletal pain intensity measured on a numerical rating scale. *Eur J Pain* 2004; 8:283–291.
19. Ostelo RW, de Vet HC, Knol DL, van den Brandt PA. 24-item Roland-Morris Disability questionnaire was preferred out of six functional status questionnaires for post-lumbar disc surgery. *J Clin Epidemiol* 2004; 57:268–276.
20. Hagg O, Fritzell P, Nordwall A. The clinical importance of changes in outcome scores after treatment for chronic low back pain. *Eur Spine J* 2003; 12:12–20.
21. Davidson M, Keating JL. A comparison of five low back disability questionnaires: Reliability and responsiveness. *Phys Ther* 2002; 82:8–24.
22. Stewart MJ, Maher CG, Refshauge KM, Herbert RD, Bogduk N, Nicholas M. Randomized controlled trial of exercise for chronic whiplash-associated disorders. *Pain* 2007; 128:59–68.
23. Kamper SJ, Ostelo RW, Knol DL, Maher CG, de Vet HC, Hancock MJ. Global Perceived Effect scales provided reliable assessments of health transition in people with musculoskeletal disorders, but ratings are strongly influenced by current status. *J Clin Epidemiol* 2010; 63:760.e1–766.e1.
24. Husted JA, Cook RJ, Farewell VT, Gladman DD. Methods for assessing responsiveness: A critical review and recommendations. *J Clin Epidemiol* 2000; 53:459–468.
25. Liang MH, Fossel AH, Larson MG. Comparisons of five health status instruments for orthopedic evaluation. *Med Care* 1990; 28:632–642.
26. Beaton DE, Hogg-Johnson S, Bombardier C. Evaluating changes in health status: Reliability and responsiveness of five generic health status measures in workers with musculoskeletal disorders. *J Clin Epidemiol* 1997; 50:79–93.
27. Garratt AM, Ruta DA, Abdalla MI, Russell IT. SF 36 health survey questionnaire: II. Responsiveness to changes in health status in four common clinical conditions. *Qual Health Care* 1994; 3:186–192.
28. Revicki D, Hays RD, Cella D, Sloan J. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. *J Clin Epidemiol* 2008; 61:102–109.
29. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. USA: Lawrence Erlbaum Associates; 1988.
30. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertain the minimal clinically important difference. *Control Clin Trials* 1989; 10:407–415.
31. 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; 63:737–745.
32. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN checklist for assessing the methodological quality of studies on measurement properties of health status measurement instruments: An international Delphi study. *Qual Life Res* 2010; 19:539–549.
33. Terwee CB, Mokkink LB, Knol DL, Ostelo RW, Bouter LM, de Vet HC. Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Qual Life Res* 2012; 21:651–657.
34. 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; 41:131–138.
35. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. *J Clin Epidemiol* 2003; 56:395–407.

Address correspondence to:

Tessa De Vrieze, PT, MT
 Department of Rehabilitation Sciences
 Catholic University of Leuven
 O&N IV Herestraat 49, Box 1510
 Leuven 3000
 Belgium

E-mail: tessa.devrieze@kuleuven.be