This is the published version of a paper published in *Neurourology and Urodynamics*.

Citation for the original published paper (version of record):

ICIQ symptom and quality of life instruments measure clinically relevant improvements in women with stress urinary incontinence.
*Neurourology and Urodynamics*, 34(8): 747-751
http://dx.doi.org/10.1002/nau.22657

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:umu:diva-99024
ICIQ Symptom and Quality of Life Instruments Measure Clinically Relevant Improvements in Women With Stress Urinary Incontinence

E. Nyström,1,2 M. Sjöström,2 H. Stenlund,1 and E. Samuelsson1

1Department of Public Health and Clinical Medicine, Umeå University, Umeå, Sweden
2Department of Public Health and Clinical Medicine, Unit of Clinical Research Center - Östersund, Umeå University, Umeå, Sweden

Aims: To determine whether changes in questionnaire scores on symptoms and condition-specific quality of life reflect clinically relevant improvements in women with stress urinary incontinence (SUI). Methods: We retrospectively analyzed questionnaires collected during a randomized controlled trial in women with SUI, that received pelvic floor muscle training (PFMT) in two different formats. We included 218 women that answered validated self-assessment questionnaires at baseline and at a 4-month follow-up. We registered changes on two questionnaires, the International Consultation on Incontinence Modular Questionnaire–Urinary Incontinence Short Form (ICIQ-UI SF) and the Lower Urinary Tract Symptoms Quality of Life (ICIQ-LUTSqol). We compared these score changes to responses from the Patient Global Impressions of Improvement (PGI-I) questionnaire. Differences were analyzed with the Spearman rho and one-way-ANOVA. The minimum important difference (MID) was the mean change in score for women that experienced a small improvement. Results: The PGI-I correlated significantly to both the ICIQ-UI SF (r = 0.547, P < 0.0001) and ICIQ-LUTSqol (r = 0.520, P < 0.0001). Thus, larger reductions in symptoms or quality of life scores were associated with greater impressions of improvement. The changes in ICIQ-UI SF and ICIQ-LUTSqol scores were significant across all PGI-I groups from “no change” to “very much improved” (P < 0.05). The MIDs were 2.52 (SD 2.56) for ICIQ-UI SF and 3.71 (SD 4.95) for ICIQ-LUTSqol. Conclusions: The change in ICIQ-UI SF and ICIQ-LUTSqol scores after PFMT reflected clinically relevant improvements in women with SUI. The MIDs established for this population may facilitate future research, treatment evaluations, and comparisons between studies. Neurourol. Urodynam. 34:747–751, 2015.

Key words: ICIQ-LUTSqol; ICIQ-UI SF; minimum important difference; pelvic floor muscle training; PGI-I; stress urinary incontinence

INTRODUCTION

Urinary incontinence is defined as any “complaint of involuntary loss of urine” by the International Continence Society and the International Urogynaecological Association. Although common among women, its prevalence varies, from 13.1% to 70.9% in different populations, and the most prevalent type is stress urinary incontinence (SUI) that is, leakage when coughing, sneezing, or upon exertion. Urinary incontinence may affect Quality of Life (QoL). The level of distress depends on the frequency, the amount of leakage, and on the subjective experience of these symptoms. For example, fear of leakage or of the associated odour might lead to social withdrawal. Therefore, the initial assessment and the evaluation of treatment should include both symptom severity and QoL. Views might differ between patients and clinicians. Thus, it is important to include reports of QoL and symptom severity directly from the patient. These reports can be collected in a standardized manner with validated self-assessment questionnaires. Of these questionnaires, many that are currently in use are highly recommended. However, the diversity of questionnaires used in research and clinical practice makes it difficult to compare studies and treatments. The International Consultation on Incontinence Modular Questionnaire (ICIQ) aims to promote a more uniform usage of questionnaires. The ICIQ has developed and included several questionnaires in a modular structure, and all are highly recommended.

Two modules of the ICIQ are the Urinary Incontinence Short Form (ICIQ-UI SF) and the Lower Urinary Tract Symptoms Quality of Life (ICIQ-LUTSqol); the former measures symptom severity, and the latter measures QoL. Both provide additive scores, and both have been validated in large patient groups that displayed various types of incontinence and LUTS. These questionnaires allow assessment of the treatment effectiveness by comparing pre- and post-treatment scores, but a given

© 2014 The Authors. Neurourology and Urodynamics published by Wiley Periodicals, Inc.
change in the overall score does not have an intuitively comprehensive meaning.8,9

The Patient Global Impression of Improvement (PGI-I) is an outcome measure that is readily understood by both the patient and the clinician, and it gives a direct reflection of the patient’s overall opinion. It is a validated scale for the patient to rate her current (treated) condition compared to her condition prior to treatment. There are seven choices, ranging from “very much better” to “very much worse.”

Correlating the ICIQ-questionnaires to the PGI-I would enable a determination of the Minimum Important Difference (MID) in the overall score. The MID is the smallest reduction in the overall score that represents an improvement detected by the patient. The MID provides important information on how a change in the ICIQ questionnaire scores can be interpreted clinically.

In previous studies on men10 and in women11 with SUI that underwent surgical treatment, the ICIQ-UI SF correlated well to the PGI-I. A correlation was also shown between the King’s Health Questionnaire (KHQ; the precursor of ICIQ-LUTSqol) and the PGI-I in women after surgical treatment for SUI.12 However, to our knowledge, no study has correlated the ICIQ-UI SF to the PGI-I in women treated with pelvic floor muscle training (PFMT), which is the recommended first-line treatment for SUI.13,14 Furthermore, we have not found any studies on the correlation between the ICIQ-LUTSqol and the PGI-I in either women or men with urinary incontinence.

Thus, in the present study, we aimed to investigate how a change in the overall scores from the ICIQ-UI SF and ICIQ-LUTSqol correlates to the PGI-I, and we aimed to establish an MID for both questionnaires in women treated with PFMT for SUI.

MATERIALS AND METHODS

Patients and Methods

We analyzed data from a recently completed randomized controlled trial (RCT), which was registered and reported at ClinicalTrials.gov (ID: NCT01032265), and described in detail by Sjöström et al.;15 that is, the additional response was considered equivalent to no effect on this aspect of quality of life (score = 1). The ICIQ-LUTSqol was derived from the validated KHQ and added to the ICIQ module, with the alteration of one question and a simplification of the scoring system. It has been proven reliable in both paper and electronic formats.16

We first calculated the overall score for both questionnaires, according to the ICIQ protocol. To save overall scores from the ICIQ-UI SF and ICIQ-LUTSqol questionnaires, as described in further detail by Sjöström et al., missing answers were replaced at baseline with the value from the 4-month follow-up, and vice versa. Six ICIQ-UISF scores and 13 ICIQ-LUTSqol scores were completed in this manner. Then, we calculated the difference between the overall scores at inclusion and after treatment. In four cases for each score, it was not possible to complete the scores at follow-up and calculate the difference. The PGI-I questionnaire consists of a single question for evaluating treatment, in our case “How is your urinary leakage now compared to before treatment?” The seven alternative responses are “Very much better,” “Much better,” “A little better,” “No change,” “A little worse,” “Much worse” and “Very much worse.” The PGI-I is validated against Incontinence Episode Frequency, stress pad test, and Incontinence Quality of Life Questionnaire.

In this study, we regarded all women as a single group, without consideration for how they were randomized in the RCT. Instead, the participants were divided into categories based on their responses to the PGI-I question. Then, the change in ICIQ-UI SF and ICIQ-LUTSqol scores between inclusion and follow-up were analyzed for each PGI-I category to determine the association.

Statistics

We calculated the Spearman rank correlation between the ICIQ-UI SF and PGI-I and between the ICIQ-LUTSqol and PGI-I. Values between 0.4 and 0.6 were considered a moderate correlation.17

No participants considered themselves “Very much worse” after treatment. Only one participant answered “Much worse;” this participant was added to the group that responded “a little worse” on the PGI-I, and together, they formed the group that we referred to as “worse.” This combined group was used in all ANOVAs, but not in the Spearman analyses.

We performed one-way ANOVAs to determine whether the mean scores at inclusion and the mean post-treatment reductions in scores on the ICIQ-UI SF and ICIQ-LUTSqol were significantly different between different PGI-I categories. We used the Welch test when there was a significant difference in homogeneity of variance between groups. Post-hoc analyses were performed according to Tukey. We also used median reductions and quartiles to illustrate the reduction within each category.

We used an anchor-based method to establish the MID; that is, we calculated a mean ICIQ score for the participants that responded “a little better” on the PGI-I, and we took that mean to reflect the smallest difference detected by the patient. The group that reported “no change” on the PGI-I had ICIQ means higher than 0, perhaps as a result of being in a study. Thus, we also calculated the difference between the ICIQ means from the

Neurourology and Urodynamics DOI 10.1002/nau
groups that responded “a little better” and “no change” on the PGI-I. This difference provided an MID for comparing one treatment to another (between-treatment MID).\textsuperscript{18,19}

All statistical analyses were performed with SPSS version 21.0 software. The groups were considered to differ significantly from each other when \( P < 0.05 \).

### Ethics

The study received ethical approval from the Regional Ethical Review Board, Umeå University (number 08–124M). All subjects provided informed consent. No reimbursements were given.

### RESULTS

At the 4-month follow-up, 218 women completed the PGI-I questionnaire and were included in this study. Four of these did not complete ICIQ-UI SF and four had no data on ICIQ-LUTSqol. The women resided all over Sweden, two-thirds in urban areas and one-third in rural areas. Of the participants, 39.0\% experienced daily leakage. Baseline data are summarized in Table I.

According to the PGI-I at follow-up, 83.0\% of participants considered their condition improved, 14.7\% unchanged, and 2.3\% deteriorated. The women were categorized into five groups according to their PGI-I response (Table II).

### ICIQ-UI SF

The mean overall score on the ICIQ-UI SF was 10.20 (SD 3.71) at baseline. These mean scores were not significantly different among the different PGI-I groups (PGI-I groups determined later; Table II).

At follow-up, the mean reduction in the overall score was 3.10 (SD 3.20). Table III summarizes the means for each PGI-I category.

We found a moderate correlation between the change in ICIQ-UI SF score and the PGI-I. The Spearman rank correlation coefficient was 0.547 (\( P < 0.001 \)); thus, the larger the reduction in score, the greater the improvement experienced by the patient. Also according to the ANOVA, the larger reduction of overall score the women had, the larger improvement they reported (\( P < 0.001 \)). There were significant differences in mean overall scores among the different PGI-I categories, except between the categories “no change” and “a little/much worse.” Figure 1 shows the distribution and the median reduction within each PGI-I category.

### TABLE I. Baseline Characteristics of Included Women (n = 218) With Stress Urinary Incontinence

<table>
<thead>
<tr>
<th>Demographics</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean age in years (SD)</td>
<td>49.24 (10.23)</td>
</tr>
<tr>
<td>Education, %</td>
<td></td>
</tr>
<tr>
<td>No higher education</td>
<td>25.2</td>
</tr>
<tr>
<td>University studies &lt; 3 years</td>
<td>21.6</td>
</tr>
<tr>
<td>University studies ≥ 3 years</td>
<td>53.2</td>
</tr>
<tr>
<td>Dwelling, %</td>
<td></td>
</tr>
<tr>
<td>Metropolitan area</td>
<td>36.3</td>
</tr>
<tr>
<td>Urban area</td>
<td>30.7</td>
</tr>
<tr>
<td>Rural area</td>
<td>26.6</td>
</tr>
<tr>
<td>Sparsely populated area</td>
<td>6.4</td>
</tr>
</tbody>
</table>

I: Patient Global Impression of Improvement; ICIQ-UI SF: International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form; ICIQ-LUTSqol: International Consultation on Incontinence Modular Questionnaire Lower Urinary Tract Symptoms Quality of Life.

### TABLE II. Mean Symptom and Quality of Life Scores at Baseline for Each Category of PGI-I

<table>
<thead>
<tr>
<th>PGI-I</th>
<th>ICIQ-UI SF</th>
<th>ICIQ-LUTSqol</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>Mean overall score (SD)</td>
<td>N</td>
</tr>
<tr>
<td>Very much better</td>
<td>16</td>
<td>10.94 (3.53)</td>
</tr>
<tr>
<td>Much better</td>
<td>54</td>
<td>10.32 (3.46)</td>
</tr>
<tr>
<td>A little better</td>
<td>108</td>
<td>10.17 (3.00)</td>
</tr>
<tr>
<td>No change</td>
<td>31</td>
<td>9.34 (2.70)</td>
</tr>
<tr>
<td>Worse</td>
<td>5</td>
<td>12.60 (4.16)</td>
</tr>
<tr>
<td>Total</td>
<td>214</td>
<td>10.20 (3.17)</td>
</tr>
</tbody>
</table>

N = the numbers of participants that answered both the PGI-I and ICIQ-UI SF or ICIQ-LUTSqol at the 4-month follow up, and hence were included in the analysis. PGI-I: Patient Global Impression of Improvement; ICIQ-UI SF: International Consultation on Incontinence Modular Questionnaire Urinary Incontinence Short Form; ICIQ-LUTSqol: International Consultation on Incontinence Modular Questionnaire Lower Urinary Tract Symptoms Quality of Life.
The MID (i.e., the mean change for the category "a little better") was 2.52 (SD 2.56). The difference between the mean changes in overall score for the groups "a little better" and "no change" was 1.58 (between-treatment MID).

ICIQ-LUTSqol

At baseline, the mean overall score for ICIQ-LUTSqol was 33.13 (SD 7.30). The baseline impact on QoL of women in the "very much better" group was significantly greater than of the group that experienced "no change" (P = 0.017), but there were no significant differences among the other groups (Table II).

At the 4-month follow-up, the mean reduction in the overall score was 4.65 (SD 6.33). There was a moderate correlation between the change in the ICIQ-LUTSqol score and the PGI-I. The Spearman rho was 0.520 (P < 0.001).

Again, the women that reported large improvements on the PGI-I questionnaire also exhibited large reductions in overall QoL score (P < 0.001). The mean change in ICIQ-LUTSqol scores differed significantly between all PGI-categories, except between the categories "no change" and "a little/much worse" (Table III). Also here, the median reductions in score displayed the same pattern as the mean reductions in score. Figure 1 shows the variation in the median reduction in ICIQ-LUTSqol score for each PGI-I category.

The MID for the ICIQ-LUTSqol was 3.71 (SD 4.95). The difference between the mean changes in overall score for the groups "a little better" and "no change" was 3.00 (between-treatment MID).

DISCUSSION

In this study, we demonstrated that the mean changes in overall scores on two questionnaires that measure symptom severity (ICIQ-UI SF) and quality of life (ICIQ-LUTSqol) correlates with the PGI-I as outcome measures. Larger reductions in the overall symptom and quality of life scores were associated with greater improvements (PGI-I) experienced after treatment. At the group level, the minimum reduction in overall score that could be detected was 2.52 for the ICIQ-UI SF and 3.71 for the ICIQ-LUTSqol.

A major strength of this study was the clinical relevance of the setting. All women had SUI, the most common type of UI, they actively sought treatment, and PFMT is the recommended first-line treatment. Other strengths were that the highly recommended ICIQ questionnaires and the PGI-I questionnaire are validated, user-friendly, and used increasingly in current research. The participants completed the questionnaires in their own homes, both at baseline and at follow-up, and there were few missing data.

One limitation was that the lack of face-to-face-treatment did not resemble the most common clinical situation today. However, there is a need for new methods of delivering non-pharmacological treatments. Studies have shown that Internet-based treatments increase access to care and might lower the barrier for seeking care. Internet-based treatments are therefore likely to increase in the future. Another possible weakness was that our study population was fairly well-educated compared to the general population. Among our participants, 74.8% had completed university studies, compared with only 42.4% in the general Swedish population of women aged 23–70 years. This may have affected our results in several ways. Higher socioeconomic status, indicated by education, is reported to yield lower satisfaction with treatment; on the other hand, neither higher nor lower educational status has been associated with higher success rates on the PGI-I. In this study, the mean changes in overall scores for the ICIQ-UI SF and ICIQ-LUTSqol correlated to the PGI-I outcome (r = 0.547 and r = 0.520) in women with SUI treated with PFMT. This was consistent with findings from Twiss et al. for the ICIQ-UI SF in men after a perineal sling operation. The correlation we found between ICIQ-LUTSqol and PGI-I was similar in strength to the correlation found by Abdel-fattah et al. between the KHQ and the PGI-I (r = 0.48) in women with SUI after surgical treatment. These moderate correlations further supported the need for assessing both symptom severity and QoL to achieve comprehensive evaluations.
We demonstrated a lower MID (2.5) for the ICIQ-UI SF than Sirls et al., who established MIDs of five and four points of reduction measured at 12 and 24 months, respectively, after a surgical treatment for SUI in women.11 These results indicate that MID changes over time. The 4-month time frame applied in our study was the same used to validate the PGI-I.8 However, MIDs might also vary, depending on the goals of the patient and of the caregiver, the pre-treatment scores, the complexity of treatment, and the anchor for the comparison.18,19 Compared to the participants of our study, the surgical cohort in the Sirls et al. study had higher preoperative scores, larger reductions after treatment,11 and most likely, different expectations of the treatment. It is reasonable to believe that women expect a better outcome from an invasive treatment with greater risk of harm than from a conservative treatment.

MIDs can be used as an aid to interpret the outcome of the questionnaires after treatment. Thus, for a treatment to be considered clinically relevant, a statistically significant post-treatment improvement in the overall score should exceed the MID. Furthermore, when comparing the effects of different treatments, a statistically significant difference cannot be considered equivalent to a clinically noticeable difference. Some researchers have argued that a between-treatment MID should be used to compare treatments. This would be calculated as the difference in scores between the categories of “a little better” and “no change.”18,19 This approach appears to be reasonable, and each researcher and clinician should consider which MID is the most appropriate for their setting. Reports directly from the patient are important in patient-centred care, particularly in a condition like SUI, where the subjective experience of symptoms affects the QoL. This study demonstrated that the symptoms and QoL scores, ICIQ-UI SF and ICIQ-LUTSqol, respectively, measured improvements that were clinically relevant in women with stress urinary incontinence treated with pelvic floor muscle training. Our results suggested that, at the group level, overall score reductions of 2.5 and 3.7, respectively, should be considered clinically relevant.

ACKNOWLEDGMENTS

We thank the participants for taking part in the RCT that enabled this study. We also thank our project coordinator, Susanne Johansson, for administration of the study. This study, like the RCT, was funded and supported by The Swedish Council for Working Life and Social Research; the Jämtland County Council; the Västerbotten County Council (ALF); and the Visare Norr, Northern County Councils, Sweden.

REFERENCES


Neurourology and Urodynamics DOI 10.1002/nau