

THE MINIMUM CLINICALLY IMPORTANT DIFFERENCE FOR THE INTERNATIONAL CONSULTATION ON INCONTINENCE QUESTIONNAIRE (ICIQ) IN WOMEN WITH STRESS URINARY INCONTINENCE

Hypothesis / aims of study

The aim of this study was to determine the minimum important difference (MID) for the International Consultation on Incontinence Questionnaire (ICIQ). The MID estimates the minimum degree of change in score correlating with patient perception of improvement.

Study design, materials and methods

Anchor-based and distribution statistical methods [1] were used to estimate the MID using data from the Trial of Midurethral Slings (TOMUS) conducted by the UITN [2].

For anchor-based analyses, Kendall's rank or Spearman correlation coefficients were first calculated to confirm that the instrument and anchors were moderately correlated ($r \geq 0.3$). Anchors included the Urogenital Distress Inventory (UDI), Incontinence Impact Questionnaire (IIQ), Incontinence episodes (IE) on 7-day bladder diary, Patient Global Impression of Improvement (PGI-I) and patient satisfaction with surgical results. For the UDI and IIQ, we compared the difference in ICIQ scores of patients with a $\geq 75\%$ reduction in IE to those with no change. Using the 7-day bladder diary, we defined the MID as the difference in ICIQ between "better" ($\geq 25\%$ decrease in IE) and no change (change in any direction between 0 and 24% in IE). The MID was the difference in ICIQ between "better" and "no change". The PGI-I is a self-rated assessment of change after treatment with 7 response levels from "very much worse" to "very much better". We used the difference in mean ICIQ between "very much better" and all others. Self-reported satisfaction with surgical results was measured as: completely satisfied, mostly satisfied, neutral, mostly dissatisfied or completely dissatisfied. We used the difference in mean ICIQ between "somewhat satisfied (completely or mostly satisfied or neutral)" and "not at all satisfied (mostly or completely dissatisfied)".

The distribution-based method was applied using effect sizes of 0.2 and 0.5 SD, small to medium effects per Cohen's rule [3]. Effect size represents individual change in terms of the number of baseline standard deviations.

Triangulation was used to examine the resulting multiple MID values in order to converge on a small range of values. The different MID estimates were graphed to visually depict the range of estimates. MID estimates and 95% CIs of each anchor- and distribution-based approach were then compared based on recommendations of Revicki et al [1].

Results

597 women were randomized in the TOMUS trial, with average age of 53 ± 11 years. 84% of subjects completed both baseline and 12-month evaluations, and 75% completed both baseline and 24-month evaluations. The ICIQ ($-10.4 @ 12$ mos., $-9.9 @ 24$ mos.), UDI ($-109.4 @ 12$ mos., $-103.9 @ 24$ mos), and IIQ ($-131.4 @ 12$ mos., $-128.4 @ 24$ mos) scores improved post-treatment at 12 and 24 months, and the number of IE from the bladder diaries declined after treatment ($-2.9 @ 12$ mos., $-2.8 @ 24$ mos.). 349 (69%) and 294 women (65%) reported being "very much better" on the PGI-I at 12 and 24 months, respectively. 460 (88%) and 403 (87%) women reported being "mostly" or "completely satisfied" with the result of bladder surgery related to the urine leakage at 12 months and 24 months, respectively.

The ICIQ met the *a priori* criteria of $r \geq 0.3$ for the anchor-based approach for all anchors (UDI, IIQ, IE, PGI-I, and Satisfaction). The mean change in ICIQ scores at 12 months and 24 months for the anchors (UDI, IIQ, IE, PGI-I, and Satisfaction) by response levels, as well as the ICIQ MID for these anchors, are presented in Table 1. MID values ranged from -4.5 to -5.7 at 12 months and from -3.1 to 4.3 at 24 months. Results of the distribution-based analyses are shown in Table 2, showing MID values that are lower than those obtained using anchor-based methods. Results of the triangulation support use of an MID for the ICIQ of -5 at 12 months and -4 points at 24 months.

Table 1. Results of Anchor-Based Methods

Anchor measure	N	12 months	N	24 months
% Change in UDI _{75%} , Mean (SD)				
Improved ($\geq 75\%$ decrease)	366	-11.7 (4.37)	289	-11.5 (4.16)
No change (0 to 75%)	137	-6.68 (4.75)	159	-7.24 (4.33)
Worse ($\geq 75\%$ increase)	1	-1.73 (---)	1	1.53 (---)
MID (95% CI) for ICIQ using Change in UDI		-5.1 (-5.9, -4.2)		-4.2 (-5.0, -3.4)
% Change in IIQ _{75%} , Mean (SD)				
Improved ($\geq 75\%$ decrease)	414	-11.4 (4.36)	356	-10.8 (4.20)
No change (0 to 75%)	80	-5.67 (5.17)	88	-6.64 (5.12)
Worse ($\geq 75\%$ increase)	5	-2.23 (4.36)	0	-----
MID (95% CI) for ICIQ using Change in UDI		-5.7 (-6.8, -4.6)		-4.2 (-5.2, -3.1)
% Change in IE _{25%} , Mean (SD)				
Improved ($\geq 25\%$ decrease)	427	-10.9 (4.69)	375	-10.2 (4.54)
No change (0 to 25%)	12	-6.11 (4.60)	12	-7.13 (4.63)
Worse ($\geq 25\%$ increase)	8	-1.04 (5.26)	14	-6.02 (4.36)

MID (95% CI) for ICIQ using %Change in IE		-4.8 (-7.5, -2.1)		-3.1 (-5.7, -0.5)
% Change in IE _{75%} , Mean (SD)				
Improved (\geq 75% decrease)	389	-11.2 (4.56)	336	-10.7 (4.40)
No change (0 to 75%)	53	-6.71 (5.23)	60	-6.51 (3.89)
Worse (\geq 75% increase)	5	-3.17 (5.51)	5	-6.99 (7.41)
MID (95% CI) for ICIQ using %Change in IE		-4.5 (-5.8, -3.1)		-4.1 (-5.3, -3.0)
PGI-I (2 categories), Mean (SD)				
Very much better (1)	336	-11.9 (4.35)	284	-11.4 (4.27)
Others (2-7)	152	-7.11 (4.87)	155	-7.21 (4.29)
MID (95% CI) for ICIQ using PGI-I		-4.8 (-5.6, -3.9)		-4.2 (-5.1, -3.4)
Satisfaction, Mean (SD)				
Somewhat satisfied = (Completely or mostly satisfied or Neutral)	454	-10.8 (4.65)	404	-10.3 (4.43)
Not at all satisfied = (Mostly or completely dissatisfied)	46	-5.64 (6.06)	38	-6.01 (5.71)
MID (95% CI) for ICIQ using Satisfaction		-5.2 (-6.6, -3.7)		-4.3 (-5.8, -2.8)

Table 2. Results of Distribution-Based Methods	
Effect Size	Δ ICIQ
0.2 SD*	-0.82
0.5 SD*	-2.05

Interpretation of results

The recommended MID for the ICIQ in a population of women with stress predominate UI is -5 for assessment at 12 months and -4 for assessment at 24 months. Surgical cohorts may overestimate the MID because of uniformly high pre-operative scores without significant variability that have a large improvement in score after treatment.

Concluding message

MID values are population-based, and an individual woman's perception may not correlate. Statistically significant changes (improvements) in ICIQ scores that meet these time-specific thresholds can be considered clinically important.

References

1. Revicki D, Hays RD, Cella D, et al. Recommended methods for determining responsiveness and minimally important differences for patient-reported outcomes. J Clin Epidemiol 2008;61(2):102–109.
2. Richter HE, Albo ME, Zyczynski HM, Kenton K, Norton PA, Sirls LT, et al. Retropubic versus Transobturator Midurethral Slings for Stress Incontinence. New Engl J Med 2010;362:2066-76.
3. Cohen J: Statistical Power Analysis for the Behavioral Sciences. 2nd edition. Hillsdale, NJ: Lawrence Erlbaum Associates; 1988.

Disclosures

Funding: Supported by cooperative agreements from the National Institute of Diabetes and Digestive and Kidney Diseases, U01 DK58225, U01 DK58229, U01 DK58234, U01 DK58231, U01 DK60379, U01 DK60380, U01 DK60393, U01 DK60395, U01 DK60397, and U01 DK60401. Support was also provided by the Eunice Kennedy Shriver National Institute of Child Health and Human Development. **Clinical Trial:** Yes **Registration Number:** U01 DK58225, U01 DK58229, U01 DK58234, U01 DK58231, U01 DK60379, U01 DK60380, U01 DK60393, U01 DK60395, U01 DK60397, and U01 DK60401. **RCT:** Yes **Subjects:** HUMAN **Ethics Committee:** Each institutions IRB **Helsinki:** Yes **Informed Consent:** Yes