2-1-2015

The Minimum Important Difference for the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form in Women with Stress Urinary Incontinence

Larry Sirls

Sharon Tennstedt

Linda Brubaker

Hae-Young Kim
New York Medical College

Ingrid Nygaard

See next page for additional authors

Recommended Citation

Sirls, L., Tennstedt, S., Brubaker, L., Kim, H., Nygaard, I., Rahn, D., Shepherd, J., & Richter, H. (2015). The Minimum Important Difference for the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form in Women with Stress Urinary Incontinence. *Neurourology and Urodynamics, 34* (2), 183-187. https://doi.org/10.1002/nau.22533

Published in final edited form as:
Neurourol Urodyn. 2015 February ; 34(2): 183–187. doi:10.1002/nau.22533.

The Minimum Important Difference for the International Consultation on Incontinence Questionnaire - Urinary Incontinence Short Form in Women with Stress Urinary Incontinence

Larry T. Sirls, MD[1], **Sharon Tennstedt, PhD**[2], **Linda Brubaker, MD**[3], **Hae-Young Kim, DrPH**[2], **Ingrid Nygaard, MD**[4], **David D. Rahn, MD**[5], **Jonathan Shepherd, MD**[6], and **Holly E. Richter, PhD, MD**[7]

[1] William Beaumont Hospital, Royal Oak, MI

[2] New England Research Institutes, Watertown, MA

[3] Loyola University Chicago, IL

[4] University of Utah, Salt Lake City

[5] University of Texas Southwestern Medical Center, Dallas

[6] University of Pittsburgh, Pittsburgh

[7] University of Alabama at Birmingham

Abstract

Introduction—Minimum important difference (MID) estimates the minimum degree of change in an instrument's score that correlates with a patient's subjective sense of improvement. We aimed to determine the MID for the International Consultation on Incontinence Questionnaire-Urinary Incontinence Short Form (ICIQ-UI SF) using both anchor based and distribution based methods derived using data from the Trial of Midurethral Slings (TOMUS).

Materials and Methods—Instruments for the anchor-based analyses included the Urogenital Distress Inventory (UDI), Incontinence Impact Questionnaire (IIQ), Patient Global Impression of improvement (PGI-I), incontinence episodes (IE) on 7-day bladder diary and satisfaction with surgical results. After confirming moderate correlation ($r \geq 0.3$) of ICIQ-UI SF and each anchor, MIDs were determined by calculating the difference between the mean instrument scores for individuals with the smallest amount of improvement and with no change. The distribution-based method of MID assessment was applied using effect sizes of 0.2 and 0.5 SD (small to medium effects). Triangulation was used to examine these multiple MID values in order to converge on a small range of values.

Results—Anchor-based MIDs range from -4.5 to -5.7 at 12 months and from -3.1 to 4.3 at 24 months. Distribution-based MID values were lower. Triangulation analysis supports a MID of -5 at 12 months and -4 at 24 months.

Corresponding Author Larry T. Sirls, MD 31157 Woodward Avenue Royal Oak, Michigan 48073 Phone 248 336 3388 Fax 248 336 3190 Lsirls@mac.com.

Conclusion—The recommended MIDs for ICIQ-UI SF are -5 at 12 months and -4 at 24 months. In surgical patients, ICIQ-UI SF score changes that meet these thresholds can be considered clinically meaningful.

Keywords

Minimum Important Difference; ICIQ-UI SF; Urinary incontinence; Quality of life; Midurethral sling

Introduction

Determination of disease-specific distress and impact on quality of life (QOL) are important parts of the assessment of patients prior to and after an intervention for stress urinary incontinence. As a part of this assessment in the Trial of Midurethral Slings (TOMUS) study[1], the International Consultation on Incontinence Questionnaire Urinary Incontinence Short Form (ICIQ-UI SF) was used[2,3]. The ICIQ-UI SF is a validated, robust measure of impact found to be sensitive to change with intervention. The minimal important difference is “the smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side-effects and excessive cost, a change in the patient's management”[4]. The smallest change in a questionnaire score that is found to correlate with meaningful symptom improvement is called the minimally important difference (MID) or, sometimes, the minimal clinically important difference (MCID). MID's are important because small numerical differences in mean questionnaire scores might be statistically significant when large sample sizes are used, but statistical significance may not be clinically significant. Knowledge of this measure helps one discern whether observed within-group or between-group *statistically* significant differences in questionnaire scores meet a threshold that is *clinically* significant or meaningful.

The ICIQ-UI SF is a brief 3-scored and 1 un-scored measure to assess the prevalence, frequency and volume of urine leakage as well as its impact on QOL. It demonstrates good construct validity and reliability and high correlation with the Sandvik Severity score[3,5]. Using outcome measures of the TOMUS study and applying both anchor-based and distribution-based methods of calculation, the purpose of this paper was to determine the MID of the ICIQ-UI SF[6]. MID may vary over time with certain chronic conditions, we evaluate MID estimates for two important time points commonly used in evaluating stress incontinence treatment outcomes. This information should help in the interpretation of data obtained from the treatment of urinary incontinence and in planning of future studies.

METHODS

Data from the Trial of Midurethral Slings (TOMUS) conducted by the Urinary Incontinence Treatment Network (UITN) were used for these analyses. The NIDDK-sponsored UITN consists of urologists and urogynecologists from nine clinical centers and a data coordinating center. TOMUS was a multi-center randomized surgical trial comparing the efficacy and morbidity of retropubic and transobturator mid-urethral sling procedures for treatment of stress urinary incontinence (SUI). Women diagnosed with predominant SUI by

self-reported symptoms for at least three months and a positive stress test at a bladder volume ≤ 300 mL and who desired surgical therapy for the treatment of SUI were eligible. Five hundred ninety seven (597) subjects were randomized in the TOMUS trial, with an average age of 53 +/- 11 years. The majority (n-473; 79%) were Non-Hispanic White, and 3% were Non-Hispanic Black. Subjects who completed both baseline and 12-month evaluations (84%) and those who completed both baseline and 24-month evaluations (75%) comprise the analytical sample.

There is no "gold standard" methodology of estimating the MID or achieving the meaningfulness of clinical trial results based on patient reported outcomes[7]. However, two commonly used methods are anchor-based methods that examine the relationship between scores on the target instrument and some independent measure, and distribution-based methods resorting to the statistical characteristics of the obtained scores[8]. We used both methods to determine MIDs for the ICIQ-UI SF.

Anchor-based MID is determined by calculating the difference between the mean instrument score for those individuals with the smallest amount of improvement and the mean instrument score of those individuals with no change[8]. For all anchor based analyses, Kendall's rank correlation coefficients or Spearman correlation coefficients were first calculated to determine whether the instrument and anchors were at least moderately correlated ($r \geq 0.3$) [4,5,7,8]. Only if this criterion was met did we proceed with the calculation of an anchor-based MID.

For this analysis, the anchors included the Urogenital Distress Inventory (UDI), Incontinence Impact Questionnaire (IIQ), Incontinence episodes (IE) on the 7-day bladder diary, Patient Global Impression of Improvement (PGI-I) and Satisfaction with surgical results.

For the UDI and IIQ anchors, we compared the difference in ICIQ-UI SF scores between patients with a ≥75% reduction to those with no change. Using criteria similar to those used to calculate the MIDs in the UITN BE-DRI Study[8], we determined the mean change in ICIQ-UI SF scores by the 7-day bladder diary as: "worse" ≥25% increase in IE; "no change" = a change in any direction between 0 and 24%; and "better" ≥25% decrease in IE. We defined the MID as the difference in ICIQ-UI SF scores between those who were "better" and those who demonstrated "no change". Since this 25% reduction was arbitrarily chosen, we also analyzed ≥75% reduction in IE on the 7-day diary[9]. The PGI-I scale is a self-rated assessment of change after treatment with 7 response levels ranging from "very much worse" to "very much better"[10]. We used the difference in mean ICIQ-UI SF scores between patients reporting "very much better" and all other responses. Satisfaction with surgical results was measured with the question: "How satisfied or dissatisfied are you with the result of bladder surgery related to the urine leakage?: --completely satisfied, mostly satisfied, neutral, mostly dissatisfied or completely dissatisfied". We used the difference in mean ICIQ-UI SF scores between patients reporting "somewhat satisfied (completely satisfied or mostly satisfied or neutral)" and those reporting "not at all satisfied (mostly dissatisfied or completely dissatisfied)".

The distribution-based method of MID assessment was applied using an effect size of 0.2 and 0.5 SD[4,7,10,11]. The effect size is calculated by dividing the difference between the baseline and follow-up scores by the standard deviation at baseline[12] and represents individual change in terms of the number of baseline standard deviations. Per Cohen's rule, a value of 0.20 is a small effect, 0.50 is a medium effect, and 0.80 is a large effect. This approach is widely understood and central to many psychometric indices[8].

The application of multiple methods to determine the MID for ICIQ-UI SF in a specific patient population will result in a range of values for the MID. Triangulation was used to examine these multiple values in order to converge on a small range of values. The different MID estimates were graphed to visually depict the range of estimates[11]. MID estimates and 95% CIs of each anchor- and distribution-based approach were then compared based on recommendations of Revicki et al.[11].

RESULTS

The ICIQ-UI SF, UDI, and IIQ scores improved post-treatment at 12 months and at 24 months. Likewise, the number of IE ascertained from the bladder diaries declined after treatment. There were 349 women (69%) and 294 women (65%) who reported that they were “very much better” on the PGII at 12 months and 24 months, respectively. Four hundred sixty (88%) and 403 (87%) women also reported that they were “Mostly satisfied” or “Completely satisfied” with the result of bladder surgery related to the urine leakage at 12 months and 24 months, respectively (Table 1).

The ICIQ-UI SF met the *a priori* criterion for using an anchor-based approach (Kendall's rank correlation coefficients or Spearman correlation coefficients (r) of ≥ 0.3) for determining the MID at 12 months and 24 months for all anchors (UDI, IIQ, IE, PGI-I, and Satisfaction with surgical results) (Table 2).

The mean change in ICIQ-UI SF scores at 12 months and 24 months for the anchors by response levels, as well as the ICIQ-UI SF MID for these anchors, are presented in Table 3. The MID (95% CI) of the changes in ICIQUI SF using UDI as the anchor varied from -5.1 (-5.9, -4.2) to -4.2 (-5.0, -3.4) at 12 months and 24 months, respectively. A similar variation was found in the MID (95% CI) of the change in ICIQ-UI SF from the IIQ: -5.7 (-6.8, -4.6) at 12 months and -4.2 (-5.2, -3.1) at 24 months. The MID (95% CI) using IE with a 25% reduction cut-point as the anchor were -4.8 (-7.5, -2.1) and -3.1 (-5.7, -0.5) at 12 months and 24 month, respectively. When the alternate cut-points were used for IE (75% reduction), the MID value was consistent with all others: -4.5 (-5.8, -3.1) at 12 months and -4.1 (-5.3, -3.0) at 24 months. The MID (95% CI) based on the PGI-I and Satisfaction measures were -4.8 (-5.6, -3.9) and -5.2 (-6.6, -3.7) at 12 months and -4.2 (-5.1, -3.4) and -4.3 (-5.8, -2.8) at 24 months, respectively.

In Table 4, distribution-based MIDs are presented for the ICIQ-UI SF using 0.2 SD, and 0.5 SD. When using the distribution-based methods, the MID values were lower than those obtained using anchor-based methods (Table 3).

Figure 1 shows the comparison of MID estimates with 95% CIs from the anchor and distribution-based approaches.

Discussion

Our study provides MID estimates for two important time points commonly recommended in stress incontinence treatment studies. Studies have suggested that MID may vary over time with certain chronic conditions[8]. The anchor-based MID at 12 months (-5) is slightly different from that at 24 months (-4) and provides a more refined threshold for investigators who will use the ICIQ-UI SF in future studies.

Researchers are increasingly incorporating patient-reported outcomes (PROs) as research outcomes. These MID values complement, but do not replace PROs as the MID is a population based observation, an individual woman's perspective may not correlate with these changes. A known limitation of anchor-based methods that rely on global ratings, particularly over longer periods of time, is that they are susceptible to recall bias and may reflect the patient's current life experience or mood[13]. Nonetheless, the FDA established guidelines for patient reported outcome measures that recommend the use of anchor-based methods to establish empiric evidence for responder data.[14]

In this analysis, we determined anchor-based methods using four patient-reported measures (UDI, IIQ, PGI-I, satisfaction) and one objective measure of severity (IE), resulting in a range of MID values. However, we recognized the clinical utility of a single value. A strength of this analysis is the use of triangulation, a systematic method to integrate the results and arrive on a single value for each of the two assessment time periods.[11,15] MIDs from clinical anchors that reflect patient-rated variables like a global rating of change are given the most weight. The global rating of change provides the single best measure of the significance of change from an individual perspective[7] and take into account more information that may affect QOL than other clinical tools.

Distribution–based methods to determine MID are based on the statistical characteristics of the obtained sample and can help support anchor-based methods or provide a MID when anchor-based methods are not available.[11] An effect size of 0.5 (change of ½ SD) is likely to be clinically significant across different patient reported questionnaires.[11] Advantages of distribution-based methods are that they provide a means for establishing change beyond some level of random variation, and provide a common metric that has equivalent meaning across measures, populations, and studies. Disadvantages of distribution-based methods are that there are few agreed-upon benchmarks for establishing clinically significant improvement and they do not in themselves provide a good sense of the clinical relevance of the change.

Klotkin showed that distribution-based and anchor-based methods were consistent with moderate impairment but were markedly different for mild or severe impairment[16]. The fact that our surgical cohort had high pre-operative scores with large post-treatment change may explain why our distribution-based MIDs were lower than the anchor-based MIDs. The mild to moderate treatment effects observed in a non-surgical SUI trial may result in lower MIDs.

However, as SUI is commonly treated surgically, we believe these values are useful for surgical researchers. Dyer showed that a non-surgical cohort of UUI patients had a higher MID for the Urogenital Distress Inventory than did a SUI surgical cohort.[8,10] This may reflect the higher QOL impact of UUI reflected in higher baseline UDI scores than in SUI patients, and that patient perception of improvement required more clinical improvement for UUI than for SUI patients.

Other strengths of this study include the prospectively collected data from a large, multi-center, surgical cohort characterized with validated clinical anchors used to establish the ICIQ-UI SF MID. Individual researchers will need to determine whether it is appropriate to generalize these MID to non-surgical or non-SUI populations.

Conclusions

The recommended MIDs for the ICIQ-UI SF in a population of women with stress predominant UI are -5 for assessment at 12 months and -4 for assessment at 24 months. Although data from surgical cohorts may overestimate MID because of uniformly high pre-operative scores without significant variability that have a large improvement in score after treatment, statistically significant changes (improvements) in ICIQ-UI SF scores that meet these time-specific thresholds can be considered clinically meaningful.

Acknowledgments

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.

References

1. Richter HE, Albo ME, Zyczynski HM, et al. Retropubic versus Transobturator Midurethral Slings for Stress Incontinence. New Engl J Med. 2010; 362:2066–76. [PubMed: 20479459]
2. Shumaker SA, Wyman JF, Uebersax JS, et al. Health-related quality of life measures for women with urinary incontinence: the Incontinence Impact Questionnaire and the Urogenital Distress Inventory Continence Program in Women (CPW) Research Group. Qual Life Res. 1995; 14(2):131–9.
3. Avery K, Donovan J, Peters TJ, et al. ICIQ: A Brief and Robust Measure for Evaluating the Symptoms and Impact of Urinary Incontinence. Neurourol Urodynam. 2004; 23(4):322–330.
4. Jaeschke R, Singer J, Guyatt GH. Ascertaining the minimal clinically important difference. Control Clin Trials. 1989; 10(4):407–15. [PubMed: 2691207]
5. Klovning A, Avery K, Sandvik H, et al. Comparison of Two Questionnaires for Assessing theSeverity of Urinary Incontinence: The ICIQ-UI SF Versus the Incontinence Severity Index Neurourol Urodyn. 2009; 28(5):411–5.
6. Crosby RD, Kolotkin RL, Williams GR. Defining clinically meaningful change in health-related quality of life. J Clin Epidemiol. 2003; 56:395–407. [PubMed: 12812812]
7. Brozek JL, Guyatt GH, Schunemann HJ. How a well-grounded minimal important difference can enhance transparency of labeling claims and improve interpretation of a patient reported outcome measure. Health and Quality of Life Outcomes. 2006; 4:69. [PubMed: 17005037]
8. Dyer KY, Xu Y, Brubaker L, et al. Minimum Important Difference for Validated Instruments in Women With Urge Incontinence. Neurourology and Urodynamics. 2011; 30(7):1319–24. [PubMed: 21563210]

9. Sirls LT, Tennstedt S, Lukacz E, et al. Condition Specific Quality of Life 24 Months After Retropubic and Transobturator Sling Surgery for Stress Urinary Incontinence. Female Pelvic Med Reconstr Surg. Sep-Oct;2012 18(5):291–5. [PubMed: 22983273]
10. Barber MD, Spino C, Janz NK, et al. The minimum important differences for the urinary scales of the Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire. Am J Obstet Gynecol. 2009; 200(5):580, e1–7. [PubMed: 19375574]
11. 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): 102–9. [PubMed: 18177782]
12. Cohen, J. Statistical Power Analysis for the Behavioral Sciences. 2nd edition.. Lawrence Erlbaum Associates; Hillsdale, NJ: 1988.
13. Metz SM, Wyrwich KW, Babu AN, et al. Validity of patient-reported helath-related quality of life global ratings of change using structural equation modeling. Qual Life Res. 2007; 16(7):1193–202. [PubMed: 17551850]
14. Food and Drug Administration. Guidance for industry patient-related outcome measures: Use in medical product development to support labeling claims. “Planning for Clinical Trial Interpretation Using a Responder Definition” (Section IV.E). :24–5.
15. Leidy NK, Wyrwich KW. Bridging the gap: using triangulation methodology to estimate minimal clinically important differences (MCIDs). COPD. 2005; 2(1):157–65. [PubMed: 17136977]
16. Kolotkin RL, Crosby RD, Williams GR. An integrated method to determine meaningful changes in health-related quality of life. J Clin Epidemiol. 2004; 57(11):1153–60. [PubMed: 15567631]

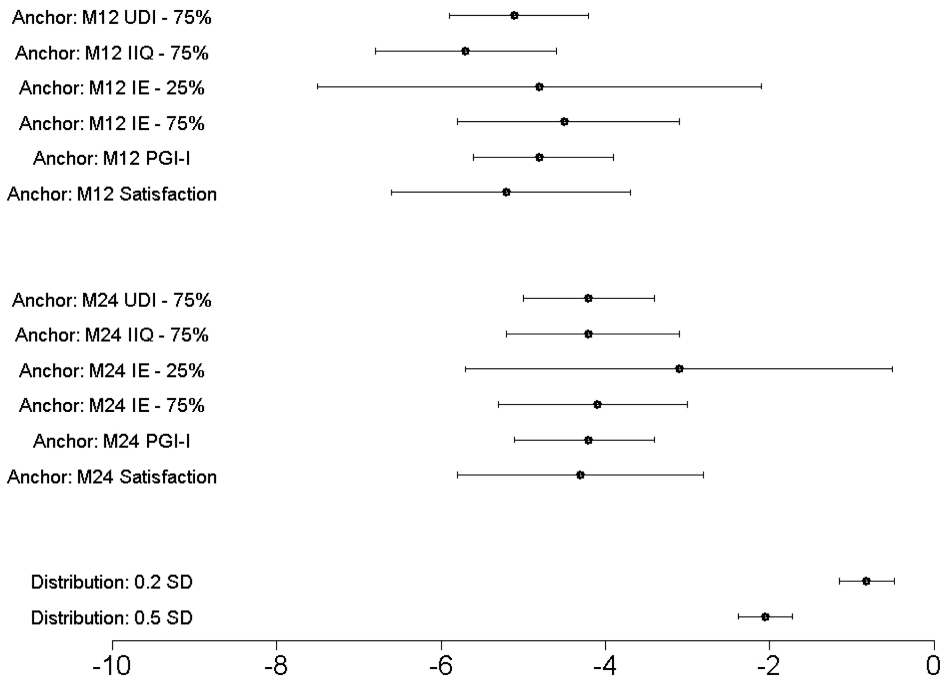

Figure 1.

Minimally important differences (MID) for the International Consultation on Incontinence Questionnaire (ICIQ) using both Anchor and Distribution-Based Methods.

Table 1

Instrument (ICIQ-UI SF) and Anchors at Baseline, 12 Months, and 24 Months

	Baseline	12 month	24 month
ICIQ-UI SF (range 0-21)	N=581 13.2 ± 4.1	N=516 2.8 ± 3.5	N=460 3.3 ± 3.5
UDI (range 0-300)	N=597 134.6 ± 45.5	N=527 25.2 ± 35.2	N=469 30.7 ± 37.3
IIQ (range 0-400)	N=597 151.5 ± 97.4	N=527 20.1 ± 47.5	N=468 23.1 ± 50.3
Incontinence episodes(IE) per day	N=593 3.3 ± 3.0	N=490 0.4 ± 1.2	N=435 0.5 ± 1.4
PGI-I			
1: Very much better	---	349 (68.7%)	294 (64.6%)
2: Much better	---	110 (21.7%)	112 (24.6%)
3: A little better	---	34 (6.7%)	31 (6.8%)
4: No change	---	7 (1.4%)	15 (3.3%)
5: A little worse	---	4 (0.8%)	1 (0.2%)
6: Much worse	---	3 (0.6%)	1 (0.2%)
7: Very much worse	---	1 (0.2%)	1 (0.2%)
Satisfaction (Urine leakage)			
Completely satisfied	---	346 (66.2%)	257 (55.6%)
Mostly satisfied	---	114 (21.8%)	146 (31.6%)
Neutral	---	15 (2.9%)	17 (3.7%)
Mostly dissatisfied	---	27 (5.2%)	23 (5.0%)
Completely dissatisfied	---	21 (4.0%)	19 (4.1%)

Table 2

Correlation between Instrument (ICIQ-UI SF) and Anchors

	ΔUDI	ΔIIQ	ΔIE	PGI-I	Satisfaction
At 12 month					
Δ ICIQ-UI SF	0.49*	0.49*	0.40*	0.37**	0.33**
At 24 month					
Δ ICIQ-UI SF	0.48*	0.44*	0.40*	0.36**	0.33**

*Correlations using Spearman rho.

**Correlations using Kendall's tau.

Table 3

Anchor-based measures and the mean change in ICIQ-UI SF by each category and the anchor-based MID (and 95% CI) at 12 months and 24 months.

Anchor measure	N	12 months	n	24 months
% Change in $UDI_{75\%}$, Mean (SD)				
Improved (≥ 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 (≥ 75% increase)	1	−1.73 (---)	1	1.53 (---)
MID (95% CI) for ICIQ-UI SF using Change in UDI**		−5.1 (−5.9, −4.2)		−4.2 (−5.0, −3.4)
% Change in $IIQ_{75\%}$, Mean (SD)				
Improved (≥ 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 (≥ 75% increase)	5	−2.23 (4.36)	0	-----
MID (95% CI) for ICIQ-UI SF using Change in IIQ**		−5.7 (−6.8, −4.6)		−4.2 (−5.2, −3.1)
% Change in $IE_{25\%}$, Mean (SD)				
Improved (≥ 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 (≥ 25% increase)	8	−1.04 (5.26)	14	−6.02 (4.36)
MID (95% CI) for ICIQ-UI SF using % Change in IE**		−4.8 (−7.5, −2.1)		−3.1 (−5.7, −0.5)
% Change in $IE_{75\%}$, Mean (SD)				
Improved (≥ 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 (≥ 75% increase)	5	−3.17 (5.51)	5	−6.99 (7.41)
MID (95% CI) for ICIQ-UI SF 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-UI SF using PGI-I†		−4.8 (−5.6, −3.9)		−4.2 (−5.1, −3.4)
Follow-up Satisfaction, Mean (SD)				
Somewhat satisfied = (Completely satisfied or Mostly satisfied or Neutral)	454	−10.8 (4.65)	404	−10.3 (4.43)
Not at all satisfied = (Mostly dissatisfied or Completely dissatisfied)	46	−5.64 (6.06)	38	−6.01 (5.71)
MID (95% CI) for ICIQ-UI SF using Satisfaction‡		−5.2 (−6.6, −3.7)		−4.3 (−5.8, −2.8)

** MID was computed based on the difference in ICIQ-UI SF mean change between “Improved” group and “No change” group.

† MID was computed based on the difference in ICIQ-UI SF mean change between “Very much better” group and “Others” group.

‡ MID was computed based on the difference in ICIQ-UI SF mean change between “Somewhat satisfied” group and “Not at all satisfied” group.

Table 4

Distribution-based MID for the ICIQ-UI SF

Effect Size	Δ ICIQ-UI SF
0.2 SD*	−0.82
0.5 SD*	−2.05

Baseline SD (=4.1) was used to compute the effect size.