The Minimum Important Differences for the Urinary Scales of the Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire

Matthew D. BARBER, MD, MHS[1], **Cathie SPINO, DSc**[2], **Nancy K. JANZ, PhD**[2], **Linda BRUBAKER, MD, MS**[3], **Ingrid NYGAARD, MD**[4], **Charles W. NAGEr, MD**[5], and **Thomas L WHEELER, MD, MSPH**[6] **for the PELVIC FLOOR DISORDERS NETWORK**

1 *Obstetrics, Gynecology, and Women's Health Institute, Cleveland Clinic, Cleveland OH*

2 *Data Coordinating Center, University of Michigan, Ann Arbor MI*

3 *Departments of Obstetrics & Gynecology and Urology, Loyola University Chicago, IL*

4 *Department of Obstetrics and Gynecology, University of Utah, Salt Lake City, UT*

5 *Department of Reproductive Medicine, University of California-San Diego, San Diego, CA*

6 *Department of Obstetrics and Gynecology, University of Alabama-Birmingham, Birmingham AL*

Abstract

Objective—Estimate the minimum important difference (MID) for the Urinary Distress Inventory (UDI), UDI-stress subscale of the Pelvic Floor Distress Inventory (PFDI), and Urinary Impact Questionnaire (UIQ) of the Pelvic Floor Impact Questionnaire (PFIQ).

Methods—We calculated MID using anchor- and distribution-based approaches from a randomized trial for non-surgical stress incontinence treatment. Anchors included a global impression of change, incontinence episodes from a urinary diary, and the Incontinence Severity Index. Effect size and standard error of measurement were the distribution methods employed.

Results—Anchor-based MIDs ranged from −22.4 to −6.4 points for the UDI, −16.5 to −4.6 points for the UDI-stress, and −17.0 to −6.5 points for the UIQ. These data were supported by two distribution-based estimates.

Conclusion—Reasonable estimates of MID are 11, 8, and 16 points for the UDI, UDI-stress subscale and UIQ, respectively. Statistically significant improvements that meet these thresholds should be considered clinically important.

Corresponding Author: Matthew D. Barber, MD, MHS, Cleveland Clinic, 9500 Euclid Ave, Desk A81, Cleveland, OH 44195, Phone: (216) 445-0439, Fax: (216) 636-5129, Email: E-mail: barberm2@ccf.org.
Reprints will not be available.

Keywords

urinary incontinence; stress urinary incontinence; health-related quality of life; minimum important difference; questionnaire; responsiveness

Introduction

Psychometrically robust instruments for measuring health-related quality of life (HRQOL) are essential to evaluate women with pelvic floor disorders.[1,2] The Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ) are two complementary condition-specific HRQOL instruments for women with urinary incontinence, fecal incontinence and pelvic organ prolapse.[3] Each has three scales: urinary, colorectal/anal, and prolapse. The PFDI and PFIQ have each been shown to be psychometrically valid, reliable, and responsive.[3–5]

The smallest change in score associated with a clinically meaningful change in a questionnaire has been called the minimum important difference (MID).[6] Because statistically significant differences may not be clinically meaningful changes, the MID of a questionnaire is essential to interpret questionnaire results when assessing within-group or between-group differences. The MID of the scales of the Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ) have been roughly estimated, but have not been specifically determined. [4]

In this analysis, we estimated the MID of the urinary scales of the PFDI and PFIQ, the Urinary Distress Inventory (UDI) and the Urinary Impact Questionnaire (UIQ), in women receiving non-surgical management of stress urinary incontinence (SUI) as part of a multi-center, randomized clinical trial. The results of this analysis should aid in interpreting and planning studies that use the PFDI and PFIQ to evaluate the impact of urinary incontinence on HRQOL.

Materials and Methods

This is an ancillary analysis of the Ambulatory Treatments for Leakage Associated with Stress (ATLAS) trial, a multi-center randomized trial conducted by the Pelvic Floor Disorders Network (PFDN) comparing behavioral therapy, incontinence pessary, and a combination of the two for treatment of stress urinary incontinence. The PFDN is sponsored by the National Institute of Child Health and Human Development and consists of seven clinical sites and a data coordinating center.

Women were eligible for the ATLAS trial if they were ≥18 years, reported stress urinary incontinence for at least 3 months, did not have advanced pelvic organ prolapse (> stage 2) and desired non-surgical treatment. Additionally, participants had predominant stress incontinence and at least 2 stress incontinence episodes on a 7-day bladder diary. Subjects were randomized to receive one of three interventions: an incontinence pessary, a 12-week behavioral therapy program, or both. The behavioral program consisted of 4 clinic visits at 2-week intervals and included pelvic floor muscle training and bladder control strategies. A detailed description of the methods of the ATLAS trial has been reported.[7]

At baseline and 3, 6 and 12 months after surgery, subjects completed a 7-day bladder diary[8] and several self-administered questionnaires to assess pelvic floor symptoms and HRQOL including the Incontinence Severity Index (ISI)[9], the PFDI and PFIQ. A global index of change was also completed at the 3, 6 and 12-month visits. ATLAS patients who completed their baseline and 3-month evaluations are the subject of this study. The three intervention groups

were combined for this analysis and treatment assignment remains blinded. Comparisons between the interventions were not performed and will be the subject of future publications.

The 46-item PFDI symptom inventory measures the degree of bother caused by a broad array of pelvic symptoms. The urinary scale, the Urinary Distress Inventory (UDI; range 0–300) has three subscales: stress, irritative, and obstructive/discomfort (range 0–100 for each). The PFIQ is a functional status measure that assesses the degree that a subject's bowel, bladder and/or pelvic symptoms impacts 31 different activities of daily living, social relationships or emotions. The urinary scale of the PFIQ is the Urinary Impact Questionnaire (UIQ; range 0–400). For the scales and subscales of the PFDI and the PFIQ, a higher score indicates worse symptom bother or poorer HRQOL.

The ISI is a two-item index that characterizes urinary incontinence severity and has demonstrated reliability, validity and responsiveness.[9–12] The index is calculated by multiplying reported incontinence frequency (five levels) by the amount of leakage (3 levels). Based on the resultant ISI value (0–12) subjects were further classified into "dry" (0), "slight" (1–2), "moderate" (3–6), "severe" (8–9) and "very severe" (12). The ISI has been validated against a 24-hour pad test and bladder diary with increasing level of severity corresponding to clinically relevant differences in these measures.[10, 12] For the global index of change, participants completed the statement "Compared to how your urinary incontinence was before treatment, do you feel you are" using a 7-point scale (very much better, much better, better, about the same, worse, much worse, very much worse).

MID can be estimated using both anchor- and distribution-based approaches and neither is known to be superior.[6] Anchor-based methods assess responsiveness in relation to an independent measure (e.g. external event or rating) to quantify the meaning of a particular degree of change in the health construct.[6, 13, 14] One way to assess the usefulness of an anchor is to determine if the anchor has at least a moderate correlation with the HRQOL measures (i.e., correlation coefficient, $r \geq .3$)[6, 13] Using Cohen's guidelines, $r = 0.1$ is small, and $r = 0.3$ is moderate.[15]

Distribution-based methods rely on the distribution of scores within a population and relate clinical significance to a change in magnitude at least equal to a statistical parameter of group data such as variability (e.g., standard deviation) or reliability (e.g., Cronbach's α). Consistent with current recommendations, we assessed the MID of the UDI, UDI-stress subscale and UIQ using both anchor-based and distribution-based techniques.[6, 14, 16]

For each approach, we evaluated the change in score from baseline to 3 months in the UDI, UDI-stress subscale and UIQ. Anchors used included the global rating of change, the ISI, and incontinence episodes (IE) on the 7-day bladder diary. Consistent with current recommendations, the MID for the global rating of change was defined as the difference between the mean change in UDI score, UDI-stress and UIQ scales reported by women who indicated they were "better" at 3 months relative to the start of treatment and those who indicated they were "about the same".[6, 16] Using criteria similar to those used to calculate the MID of the Incontinence Quality of Life (I-QOL)[17], we determined the mean change in UDI, UDI-stress subscale and UIQ scores that corresponded to three outcomes assessed by the bladder diary: "worse" ≥ 25% increase in number of IE; "no change" = a change in any direction between 0 and 24%; and "better" ≥ 25% decrease in the number of IE. We defined the MID as the difference in scores between those who were "better" and those who demonstrated "no change." Total incontinence episodes per week, regardless of type, were used for the MID determination for this anchor for each of the scales. For the UDI-stress scale, we also investigated whether considering only the stress incontinence episodes recorded on the bladder diary impacted this MID estimate. The MID using the ISI was defined as an improvement of

one severity level from baseline to the 3 month evaluation (e.g. Severe to moderate, slight to dry, etc.) and determined the mean change in score of the UDI, UDI-stress subscale and UIQ that corresponded to each of these transitions. We defined the MID using the ISI as the anchor as the difference in change in score between those subjects who had a one level improvement on ISI and those whose ISI category did not change from baseline.

Distribution-based measures of MID included effect size and standard error of measurement (SEM).[6] The effect size was based on the baseline standard deviation of each outcome; MID was calculated as 0.5 SD (medium effect size) and 0.2 SD (small effect size).[15] SEM was calculated as baseline standard deviation multiplied by the square root of (1-Cronbach's α) for each scale under consideration; 1 SEM was considered an estimate of MID.[18] 95% confidence intervals (CI) of the various MID estimates were calculated. For anchor-based methods, the CI was based on the standard 95% CI for difference between two population means; for distribution-based methods, the CI was based on the standard 95% CI for the population variance (effect size method)[19] or the bootstrap CI (SEM).[20]

MID estimates and 95% CIs of each anchor- and distribution-based approach was compared and recommendations for the MID of the UDI, UDI-stress subscale and UIQ were made by consensus consistent with the recommendations of Revicki et al.[16]

Results

Of the 445 subjects who were randomized in the ATLAS trial, about 75% of subjects completed both the baseline and 3-month evaluations and are the subject of this analysis (76% for UDI, 75% for UDI-Stress and 74% for UIQ). Baseline demographics are listed in Table 1. Three months after initiating treatment, significant improvements in UDI, UDI-stress and UIQ scores were noted (Table 2). Similarly, there were significant improvements in IE on the 7-day diary and ISI scores. Mean (SD) changes from baseline to 3 months after treatment in UDI, UDI-stress and UIQ scores were −34 (39), −19 (22), and −34 (44) points, respectively. The correlations between all anchor-based measures and HRQOL measures exceeded the recommended criteria of r ≥ .3.

Table 3 presents the change in UDI, UDI-stress and UIQ by global rating of change response category. The MID (95% CI) based on the global rating was −6.4 (−19.4, 6.5), −4.6 (−12.7, 3.5) and −6.5 (−22.8, 9.8) for the UDI, UDI-stress and UIQ, respectively. MID determined by the other anchor-based methods (ISI and bladder diary) were greater than those for the global rating and similar to the distribution-based findings.

Seventy-one percent of subjects satisfactorily completed the 7-day bladder diary at baseline and at 3 months. Of this group, 86% demonstrated a 25% or greater decrease in number of IE on diary from baseline to 3 months meeting our predefined criteria of "Improved". (Table 4) The MID (95% CI) based on the bladder diary were −22.4 (−36.5, −8.2), −16.5 (−24.5, −8.3), and −17.0 (−32.9, −1.1) for the UDI, UDI-stress and UIQ, respectively. We explored other potential cut-points in improvement in incontinence episode frequency for differentiating "Improved" from "No change." Regardless of dichotomous cut-point chosen, MID estimates were not substantially different from those of our predefined cut-point of 25%. For example, if subjects with 50% or greater improvement in incontinence episode frequency are considered "Improved" while those with less than 50% improvement were categorized as "No change," the MID values are −20.9, −16.8 and −22.8 for the UDI, UDI-stress and UIQ, respectively. Considering only stress incontinence episodes recorded on the bladder diary for the UDI-stress scale, did not significantly change MID estimate for this scale (data not shown).

The relationship between change in the urinary scales of the PFDI and PFIQ and changes in incontinence severity category as measured by the ISI is shown in Table 5. For each increasing

level of improvement in ISI severity category, there is a stepwise improvement in UDI, UDI-stress and UIQ. The MID (95% CI) using ISI as the anchor were −11.1 (−19.8, −2.3), −7.5 (−12.7, −2.3), and −16.0 (−26.4, −5.7) for the UDI, UDI-stress and UIQ, respectively.

MID estimates (95% CI) based on the distribution-based criteria for the UDI, UDI-stress subscale and UIQ are as follows: 0.5 SD corresponds to an improvement in score of −20.5 (−18.8, −21.9), −9.8 (−9..1, −10.6), and −28.7 (−26.7, −31.1) points, while 0.2 SD corresponds to score improvement of −8.1 (−8.8, −7.5), −3.9 (−4.2, −3.6), and −11.5 (−12.4, −10.7) points and 1 SEM corresponds to −15.3 (−14.2, −16.4), −13.1 (−12.3, −13.9), and −11.7 (−10.9, −12.6) points, respectively. Figure 1 shows the comparison of MID estimates with 95% CIs from the anchor- and distribution-based approaches.

Discussion

This study used three anchor-based approaches and two distribution-based approaches to establish the MID for the urinary scales of the PFDI and PFIQ. The anchor-based approaches included one measure of the patient's perspective (global rating of change) and two clinical measures of incontinence severity (ISI and number of incontinence episodes recorded on the bladder diary). Using anchor-based methods, the range of MID for the UDI was −22.4 to −6.4 points, for the UIQ was −16.5 to −4.6 points, and for the UDI-stress subscale was −17.0 to −6.5 points. These data were supported by two distribution-based estimates, effect size (1/2 SD) and 1 SEM.

When multiple approaches are used to determine the MID of a scale, a range of values rather than a single point estimate are expected, as was seen in this study. Clinically, a more narrow range or even a single point of MID for each urinary scale would be more helpful than the somewhat broad range determined by our calculations. It is recommended that a single MID value or narrow range be selected for a given scale by integrating the results from the multiple approaches in a systematic way, sometimes called triangulation.[16, 21] Using this method, MID derived from anchor-based methods that reflect patient-rated and disease-specific variables are given the most weight.[16] The global rating of change, the most commonly used anchor for MID determination, represents the best measure of the significance of change from an individual perspective.[6].

For the scales considered in this study, the MID estimates from the patient's perspective are considerably smaller than when clinical criteria are used. When determining the MID, it is recommended that the patient's perspective be given the most weight, although the clinician's perspective should also be considered.[16] Retrospective self-reports like a global rating of change are subject to recall bias and have a tendency to reflect a subject's current health state more than a change from baseline.[14]

Therefore, when examining the MID estimates and 95% CIs from all anchor-based approaches and considering the potential for some recall bias on our global health rating, we propose MID thresholds of −11, −8, and −16 for the UDI, UDI-stress subscale and UIQ respectively. These values represent the MID determined by a clinically relevant anchor, a one level change in the ISI and are higher than the MID estimates obtained from the global rating of change. Thus, values that meet or exceed these thresholds represent the minimal scores that can be considered clinically relevant to both the patient and the clinician. We recognize, based on MID estimates of the global rating of change, that smaller scores may be clinically relevant to some patients, however it is unclear if these smaller scores would be relevant from a physician's perspective. MID estimates obtained from the bladder diary represent more conservative estimates of clinically meaningful change than the ISI or global rating of change, but do not represent the least or minimum change clinically relevant to patient's and clinicians in our opinion.

In MID determination, distribution-based methods can support and help interpret estimates from anchor-based methods and be used in situations where anchor-based approaches are unavailable.[16] There are increasing data and growing consensus that an effect size of 0.5 (or change of ½ SD) is a conservative estimate that is likely to be clinically significant across different patient-reported questionnaires.[16, 22] Sloan et al proposed that, in the absence of other information, ½ SD is a reasonable and scientifically supportable estimate of meaningful effect.[22] The MID for the scales of the PFDI and PFIQ were previously estimated using this approach in a cohort of patients undergoing surgery for pelvic organ prolapse with and without SUI [4] Our study improves on this because our participants were not undergoing surgery (a treatment expected to show a large difference and thus overestimate the MID) [6] and all had urinary incontinence and therefore using measures that assess incontinence was more appropriate. Trials of non-surgical SUI therapies, such as ATLAS are particularly suited for MID determination because mild to moderate treatment effects are more common. Not surprisingly MID estimates corresponding to ½ SD were considerably less in our study than those derived from the surgical cohort (−30 versus − 20 points, and −49 versus − 29 points for the UDI and UIQ respectively).

Although the ½ SD approach provides scores that are certainly clinically significant and meaningful, they are not necessarily minimal. Consistent with other studies, MID estimates corresponding to ½ SD in our study represent the upper boundary of the range of MID estimates identified.[16] There is also increasing evidence that a change in score equivalent to 1 SEM is a valid alternative for estimating MID in patient-reported health outcome measures.[18] In this study, the 1 SEM criteria corresponded to MID estimates within the range of the anchor-based approaches for the UDI, UDI-stress subscale and UIQ.

As the method of scoring and range of possible scores differ between instruments, it is not meaningful to directly compare MID across different measures. However, one can compare the percentage of the total scale score represented by the MID (i.e. dividing the MID by the highest possible scale score).[23] The range of anchor-based MID estimates seen in this study represents 2.1 to 7.5% change of the total score for the UDI, 1.6% to 4.3% for the UIQ and 4.6 to 16.5% for the UDI-Stress subscale. These are consistent with results of several other HRQOL questionnaires including the Medical Outcomes Study Short-form 36 (SF-36) (range 3 to 6%)[24], the Functional Assessment of Chronic Illness Therapy (FACIT) scales (3.7 to 12.5%)[25], the Incontinence Quality of Life (I-QOL) scale (2.5%)[26] and Kings Health Questionnaire scales (5%).[27]

The strengths of this study include the use of multiple approaches to triangulate MID estimates following currently recommended guidelines, the use of validated and widely accepted SUI outcome measures as anchors, the large sample size, and the wide breadth of treatment response allowing an assessment of minimal change. A limitation is that we only determined MID for the urinary scales of the PFDI and PFIQ; MID estimates for the colorectal anal and prolapse scales will require further study using different patient populations. Until such estimates are available, a change of ½ SD or greater seems a reasonable conservative estimate for a clinically important change. An additional limitation of the study is that there were too few subjects who had a decline in urinary function to provide estimates for an MID for deterioration. The MID values proposed from this study represent MID for improvement only. It is also worth noting that MID estimates can vary across populations.[16]

In conclusion, MID can help researchers and clinicians understand whether HRQOL score differences between treatment groups or if changes within one group over time are clinically meaningful.[23] From this study we recommend that a difference of 11, 8, and 16 points can be considered reasonable estimates of MID for the UDI, UDI-stress subscale and UIQ, respectively. Statistically significant improvements that meet or exceed these thresholds should

be considered clinically important. However, some patients with changes in scores less than these estimates may perceive clinical important improvements. MID values should be confirmed based on accumulating evidence from multiple studies and, with increasing evidence, we will become more precise in the MID values for the urinary scales of the PFDI and PFIQ.

Acknowledgements

References

1. DONOVAN, J.; BOSCH, R.; GOTOH, M., et al. Symptom and Quality of Life Assessment. In: Abrams, P.; Cordozo, L.; Koury, S.; Wein, A., editors. 3rd International Consultation on Incontinence. Paris: Health Publication Ltd; 2005.

2. WEBER AM, ABRAMS P, BRUBAKER L, et al. The standardization of terminology for researchers in female pelvic floor disorders. Int Urogynecol J Pelvic Floor Dysfunct 2001;12:178–86. [PubMed: 11451006]

3. BARBER MD, KUCHIBHATLA MN, PIEPER CF, BUMP RC. Psychometric evaluation of two comprehensive condition-specific quality of life instruments for women with pelvic floor disorders. Am J Obstet Gynecol. 2001in press

4. BARBER MD, WALTERS MD, CUNDIFF GW. Responsiveness of the Pelvic Floor Distress Inventory (PFDI) and Pelvic Floor Impact Questionnaire (PFIQ) in women undergoing vaginal surgery and pessary treatment for pelvic organ prolapse. Am J Obstet Gynecol 2006;194:1492–8. [PubMed: 16647933]

5. WREN PA, JANZ NK, BRUBAKER L, et al. Reliability of health-related quality-of-life measures 1 year after surgical procedures for pelvic floor disorders. Am J Obstet Gynecol 2005;192:780–8. [PubMed: 15746672]

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. RICHTER HE, BURGIO KL, GOODE PS, et al. Non-surgical management of stress urinary incontinence: ambulatory treatments for leakage associated with stress (ATLAS) trial. Clin Trials 2007;4:92–101. [PubMed: 17327249]

8. LOCHER JL, GOODE PS, ROTH DL, WORRELL RL, BURGIO KL. Reliability assessment of the bladder diary for urinary incontinence in older women. J Gerontol A Biol Sci Med Sci 2001;56:M32–5. [PubMed: 11193230]

9. SANDVIK H, SEIM A, VANVIK A, HUNSKAAR S. A severity index for epidemiological surveys of female urinary incontinence: comparison with 48-hour pad-weighing tests. Neurourol Urodyn 2000;19:137–45. [PubMed: 10679830]

10. SANDVIK H, ESPUNA M, HUNSKAAR S. Validity of the incontinence severity index: comparison with pad-weighing tests. Int Urogynecol J Pelvic Floor Dysfunct 2006;17:520–4. [PubMed: 16547687]

11. SANDVIK H, HUNSKAAR S, SEIM A, HERMSTAD R, VANVIK A, BRATT H. Validation of a severity index in female urinary incontinence and its implementation in an epidemiological survey. J Epidemiol Community Health 1993;47:497–9. [PubMed: 8120507]

12. HANLEY J, CAPEWELL A, HAGEN S. Validity study of the severity index, a simple measure of urinary incontinence in women. Bmj 2001;322:1096–7. [PubMed: 11337439]

13. GUYATT GH, OSOBA D, WU AW, WYRWICH KW, NORMAN GR. Methods to explain the clinical significance of health status measures. Mayo Clin Proc 2002;77:371–83. [PubMed: 11936935]

14. WYRWICH KW, BULLINGER M, AARONSON N, HAYS RD, PATRICK DL, SYMONDS T. Estimating clinically significant differences in quality of life outcomes. Qual Life Res 2005;14:285–95. [PubMed: 15892420]

15. COHEN, J. Statistical power analysis for the behavioral sciences. New York: Acacemic Press Inc; 1977.

16. 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;61:102–9. [PubMed: 18177782]

17. PATRICK DL, MARTIN ML, BUSHNELL DM, YALCIN I, WAGNER TH, BUESCHING DP. Quality of life of women with urinary incontinence: further development of the incontinence quality of life instrument (I-QOL). Urology 1999;53:71–6. [PubMed: 9886591]

18. REJAS J, PARDO A, RUIZ MA. Standard error of measurement as a valid alternative to minimally important difference for evaluating the magnitude of changes in patient-reported outcomes measures. J Clin Epidemiol 2008;61:350–6. [PubMed: 18313559]

19. OTT, L. An introduction to statistical methods and data analysis. Belmont, CA: Wadsworth Publishing Company, Inc; 1977.

20. EFRON, B.; TIBSHIRANI, R. An Introduction to the Bootstrap. New York: Chapman & Hall/CRC; 1993.

21. LEIDY NK, WYRWICH KW. Bridging the gap: using triangulation methodology to estimate minimal clinically important differences (MCIDs). Copd 2005;2:157–65. [PubMed: 17136977]

22. SLOAN JA, CELLA D, HAYS RD. Clinical significance of patient-reported questionnaire data: another step toward consensus. J Clin Epidemiol 2005;58:1217–9. [PubMed: 16291464]

23. YOST KJ, ETON DT. Combining distribution- and anchor-based approaches to determine minimally important differences: the FACIT experience. Eval Health Prof 2005;28:172–91. [PubMed: 15851772]

24. WARE, JE, JR. SF-36 Health Survey: Manual and Interpretation Guide. Boston: Heath Institute, New England Medical Center; 1993.

25. YOST KJ, CELLA D, CHAWLA A, et al. Minimally important differences were estimated for the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) instrument using a combination of distribution- and anchor-based approaches. J Clin Epidemiol 2005;58:1241–51. [PubMed: 16291468]

26. YALCIN I, PATRICK DL, SUMMERS K, KINCHEN K, BUMP RC. Minimal clinically important differences in Incontinence Quality-of-Life scores in stress urinary incontinence. Urology 2006;67:1304–8. [PubMed: 16750246]

27. KELLEHER CJ, PLEIL AM, REESE PR, BURGESS SM, BRODISH PH. How much is enough and who says so? Bjog 2004;111:605–12. [PubMed: 15198790]

Appendix - Pelvic Floor Disorders Network Contributors

Cleveland Clinic

Mathew D. Barber, MD, MHS, Principal Investigator

Marie Fidela R. Paraiso, MD, Co-Investigator

Mark D. Walters, MD, Co-Investigator

J. Eric Jelovsek, MD, Co-Investigator

Firouz Daneshgari, Co-Investigator

Linda McElrath, RN, Research Nurse Coordinator

Donel Murphy, RN, MSN, Research Nurse

Cheryl Williams, Research Assistant

Duke University

Anthony G. Visco, MD, Principal Investigator

Jennifer Wu, MD, Co-Investigator

Alison Weidner, MD, Co-Investigator

Cindy Amundsen, MD, Co-Investigator

Mary J. Loomis, RN, BSN, Research Coordinator

Loyola University, Chicago

Linda Brubaker, MD, MS, Principal Investigator

Kimberly Kenton, MD, MS, Investigator

MaryPat FitzGerald, MD, MS, Investigator

Elizabeth Mueller, MD, MSME, Investigator

Kathy Marchese, RN, Study Coordinator

Mary Tulke, RN, Study Coordinator

University of Alabama at Birmingham

Holly E. Richter, PhD, MD, Principal Investigator

R. Edward Varner, MD, Co-Investigator

Robert L. Holley, MD, Co-Investigator

Thomas L. Wheeler, MD, Co-Investigator

Patricia S. Goode, MD, Co-Investigator

L. Keith Lloyd, MD, Co-Investigator

Alayne D. Markland, DO, Co-Investigator

Velria Willis, RN, BSN, Research Coordinator

Nancy Saxon, BSN, Research Nurse Clinician

LaChele Ward, LPN, Research Specialist

Lisa S. Pair, CRNP

University of Michigan

Morton B. Brown, PhD, Co-Investigator

Cathie Spino, PhD, Principal Investigator

John T. Wei, MD, MS, Co-Principal Investigator

Beverly Marchant, RN, BS, Project Manager

Donna DiFranco, BS, Clinical Monitor

John O.L. DeLancey, MD, Co-Investigator

Dee Fenner, MD, Co-Investigator

Nancy K. Janz, PhD, Co-Investigator

Wen Ye, PhD, Statistician

Zhen Chen, MS, Statistician

Yang Wang Casher, MS, Database Programmer

University of Texas, Southwestern

Joseph Schaffer MD – Principal Investigator

Clifford Wai, MD - Co-Investigator

Marlene Corton, MD - Co-Investigator

Gary Lemack, MD - Co-Investigator

Kelly Moore - Research Coordinator

David Rahn, MD

Amanda White, MD

Shanna Atnip, NP

Margaret Hull, NP

Pam Martinez, NP

Deborah Lawson, NP

University of Utah

Ingrid Nygaard, MD, Principal Investigator

Peggy Norton, MD, Co-Investigator

Linda Freeman, RN, Research Coordinator
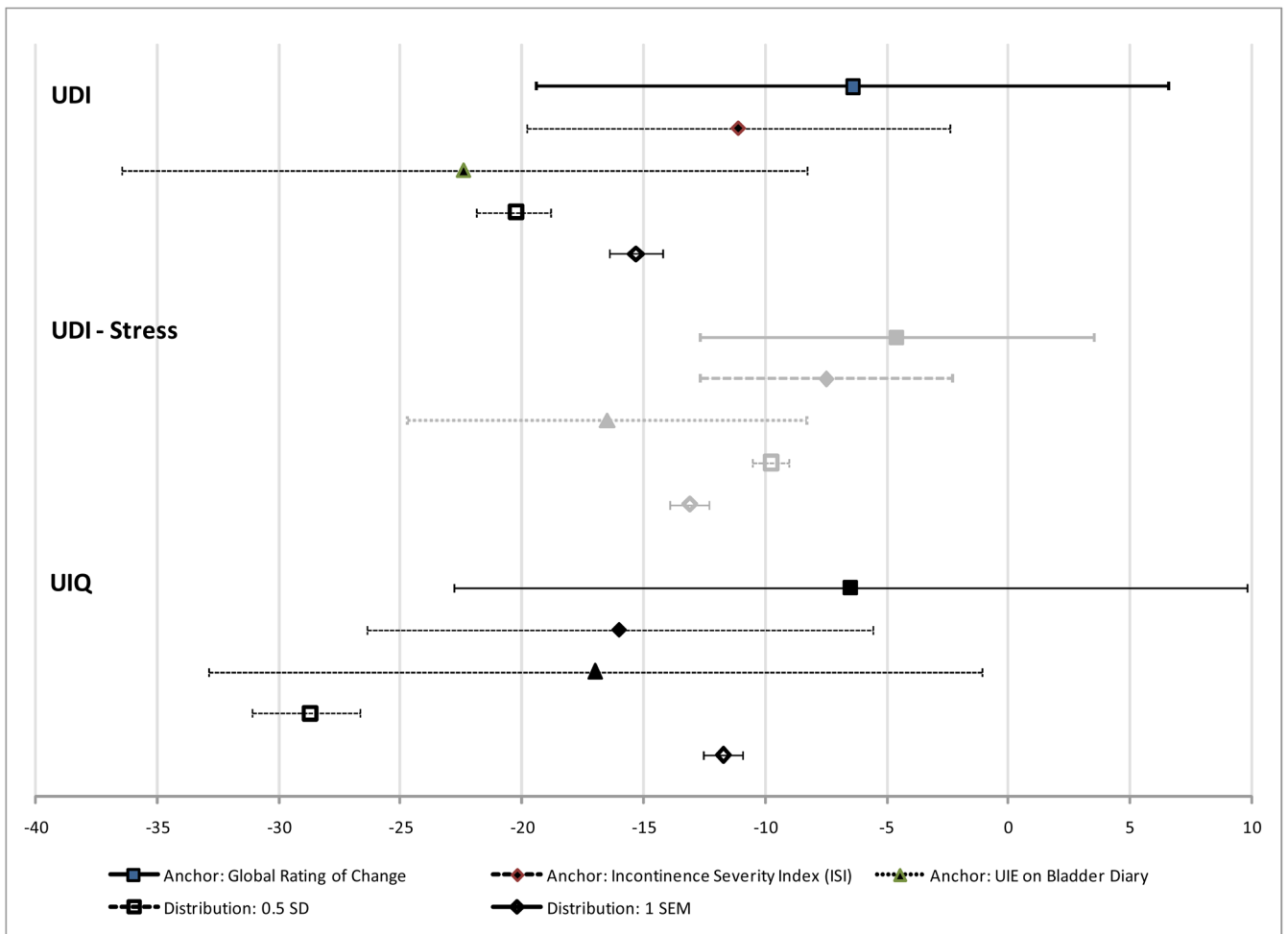
NIH Project Scientist

Susan Meikle, MD

Figure 1.
Minimally important differences (MID) for Urinary Distress Inventory (UDI), Urinary Distress
Inventory – Stress subscale(UDI-Stress) and Urinary Impact Questionnaire (UIQ) for Anchor-
and Distribution-Based Methods.

Table 1

Demographics

N	444
Mean age, years (SD)	50 (12)
Median parity (range)	2 (0–11)
Mean BMI, kg/m^2 (SD)	29 (7)
Race (%):	
White	85
Black	10
Asian	2
Other	3
Hispanic ethnicity (%)	7
Highest education (%)	
Less than high school	4
High school	19
Some college	33
College graduate	24
Graduate or professional degree	20

Table 2

Change in UDI, UIQ, Incontinence episodes on bladder diary, and Incontinence severity index

	N	Baseline	3 months	Δ (3 months − baseline)	% Δ [*]	P[#]
UDI	337	80±40	46±36	−34±39	−40 (55)	<0.01
UDI – stress	335	47±19	28±21	−19±22	−37 (62)	<0.01
UIQ	327	71±57	37±45	−34±44	−47 (85)	<0.01
Incontinence episodes on bladder diary (7 days)	316	17±20	5±8	−12±20	−86 (41)	<0.01
Incontinence Severity Index [**]:						<0.01
Dry (0)		0 (0)	4 (1)			
Slight (1–2)		18 (4)	145 (44)			
Moderate (3–6)		254 (59)	143 (43)			
Severe (8–9)		120 (28)	29 (9)			
Very Severe (12)		36 (8)	8 (2)			
Total		428	329			

Data expressed as mean ± SD unless otherwise noted.

[*] Data presented as median (interquartile range)

[#] P-values from one-sample test of change.

[**] Data presented as n (%)

Table 3

Patient global impression of improvement (PGI-I) and change in UDI & UIQ

	n	UDI	UDI-stress	UIQ
Very much better	84	−51±38	−29±23	−45±42
Much better	118	−40±39	−23±20	−42±44
Better	103	−20±32	−11±20	−22±42
About the same	32	−13±31	−7±19	−16±31
Worse#	5	−19±37	0±23	+22±82
MID*	135	−6.4	−4.6	−6.5
95% CI		−19.4, 6.5	−12.7, 3.5	−22.8, 9.8

Data presented as mean change in score ±SD from baseline to 3 months #Includes subjects who indicated they were "worse," "much worse" or "very much worse"

*
Difference in "Better" and "About the same" category means.

Table 4

Change in number of incontinence episodes on bladder diary and change in UDI & UIQ

	Definition	n	UDI	UDI-Stress	UIQ
Improved	25% or greater decrease in number of IE on diary	273	−38±39	−22±23	−38±42
No change	Change in IE on diary less than ±25%	32	−16±25	−6±15	−21±47
Worse	25% or greater increase in number of IE on diary	11	−8±31	+1±14	+9±48
MID*		305	−22.4	−16.5	−17.0
95% CI			−36.5, −8.2	−24.5, −8.3	−32.9, −1.1

Data presented as mean change in score±SD

*
Difference in "Improved" and "No change" category means

Table 5

Change in Severity Index levels and change in UDI & UIQ

Change in Severity level from baseline to 3 months	n	UDI	UDI-Stress	UIQ
Very severe to severe	6	−24±30	−19±20	−42±60
Severe to moderate	48	−30±30	−16±19	−39±42
Moderate to slight	99	−40±41	−24±23	−38±45
Slight to dry	0			
Very Severe to Moderate	13	−45±41	−16±21	−53±40
Severe to slight	27	−54±27	−34±18	−59±42
Moderate to dry	2	+22±36	+15±26	+4+6
No change	107	−25±32	−14±20	−22±35
One level improvement	153	−36±38	−21±22	−39±45
Two level improvement	42	−47±35	−26±23	−54±42
MID[*]	260	−11.1	−7.5	−16.0
95% CI		−19.8, −2.3	−12.7, −2.3	−26.4, −5.7

Data presented as mean change in score±SD

[*] difference between One level improvement and No change category mean