



Published in final edited form as:

*Neurogastroenterol Motil.* 2016 August ; 28(8): 1148–1156. doi:10.1111/nmo.12809.

## Validating Endpoints for Therapeutic Trials in Fecal Incontinence

Jessica Noelting, MD<sup>1</sup>, Alan R. Zinsmeister, PhD<sup>2</sup>, and Adil E. Bharucha, MBBS, MD<sup>3</sup>

<sup>1</sup> Division of Gastroenterology and Hepatology, Department of Internal Medicine, Rochester, MN. Dr. Noelting is now with the Division of Gastroenterology, Mayo Clinic, Scottsdale, AZ 85259

<sup>2</sup> Division of Biomedical Statistics and Informatics, Department of Health Sciences Research, Mayo Clinic, Rochester, MN 55905

<sup>3</sup> Division of Gastroenterology and Hepatology, Department of Internal Medicine, Rochester, MN 55905

### Abstract

**Background**—A 50% or greater reduction in the frequency of fecal incontinence (FI) recorded with daily bowel diaries is the primary endpoint in clinical trials of FI. Whether this difference is clinically important is unknown. The relationship between FI symptoms recorded with daily and weekly instruments is unknown. The contribution of psychological factors to quality of life (QOL) in FI is unclear.

**Methods**—FI severity was assessed with daily bowel diaries and periodic questionnaires (Fecal Incontinence Severity score [FISS], FIQOL, 36-Item Short Form Health Survey [SF-36], and Hospital Anxiety and Depression scales) for 4 weeks before and during double-blind randomization to placebo or clonidine in 44 women with FI. The reduction in FI frequency was compared to the minimal clinically-important difference (MCID) computed from the FISS. Endpoints of FI were compared between daily and weekly diaries.

**Key Results**—The FISS exceeded the MCID in 75% and 83% of patients in whom the FI frequency declined by 50%-74% and 75% respectively. Parameters of FI measured with daily and weekly instruments were significantly correlated. The daily parameters explained 71% of the inter-patient variation in the FISS. The SF-36 health scores, rather than the FISS rating, explained a majority of the inter-subject variation in FIQOL.

**Conclusions & Inferences**—Most patients who report a 50% reduction in FI frequency experience a clinically-important improvement. Weekly questionnaires accurately assess the severity of FI. Self-reported physical and mental health explained a greater proportion of the variance in FIQOL than FI symptom severity.

---

**Corresponding Author:** Adil E. Bharucha, MD, Division of Gastroenterology and Hepatology, Mayo Clinic, 200 First St SW, Rochester, MN 55905. bharucha.adil@mayo.edu.

**Specific author contributions:** Jessica Noelting - Analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; approved the final draft submitted

Alan R. Zinsmeister - Statistical analysis; approved the final draft submitted

Adil E. Bharucha - Study concept and design; analysis and interpretation of data; drafting of the manuscript; critical revision of the manuscript for important intellectual content; approved the final draft submitted

**Disclosures:** No competing interests declared.

## Keywords

bowel diaries; endpoints; fecal incontinence; quality of life

Fecal incontinence (FI) is a common condition that affects between 7% and 15% in adult women and men (1-5) and can significantly impair the quality of life (QOL) (6). A Cochrane review observed that there was limited evidence to guide clinicians in selecting therapies for FI, which mostly focus on treating diarrhea rather than FI (7). Moreover, “all (drug) trials were of short duration” and none “provided long-term results of treatment.” The response to sacral nerve stimulation (SNS) is sustained in a substantial proportion of patients 5 years after the procedure (8).

When maintained reliably, daily bowel diaries are the gold standard for recording bowel habits and FI in the short-term because they avoid recall bias and provide insights into the relationship between bowel patterns (e.g., stool consistency) and FI (9). However, daily diaries are burdensome to complete, hence perhaps not as convenient for evaluating long-term efficacy (10). In the SNS trial, long-term data were obtained with 2-week bowel diaries and questionnaires evaluating the severity of FI and its impact on QOL at every annual follow-up visit (8). However, the correlation between severity of FI assessed with daily bowel diaries and questionnaire-based endpoints has not been evaluated.

Clinical trials regard a 50% or greater reduction in the number of days or episodes of FI recorded with bowel diaries as the primary endpoint. However, it has been suggested that a 50% reduction does “not constitute a clinically meaningful improvement from the patient's perspective, mainly because the degree of improvement does little to allay the anxiety of the patient that an accident may occur at any time” (11). Indeed, it is necessary to determine whether a statistically significant change in the frequency of FI is also clinically significant. The minimum clinically important difference (MCID), which is the smallest change detected by an instrument that is associated with a clinically meaningful change (12), has been used to identify clinically meaningful changes in several conditions, including urinary incontinence (13). The MCID can be estimated using anchor- and distribution-based approaches (12). Anchor-based methods assess responsiveness in relation to an independent measure (e.g., another rating) (12). Because there is no universally accepted anchor in FI, the present study relied upon distribution-based methods as described later.

While SNS improves the symptoms and QOL in FI, only 5%-15% of the variance in FIQOL can be explained by the improvement in symptoms (14). FI is associated with anxiety and depression (15), which may impact the QOL (16). Thus, it is conceivable that psychological distress partly explains the effect of therapy on FIQOL.

During a randomized, controlled trial of clonidine for FI, participants recorded their bowel symptoms in daily bowel diaries and weekly questionnaires (17). This study sought to assess (i) FI endpoints evaluated with daily and weekly instruments; (ii) whether treatment-related differences in FI frequency are clinically significant; (iii) FI symptom severity recorded with two instruments (Fecal Incontinence Severity Index [FISI] and Fecal Incontinence Severity Score [FISS]) and separately FI symptom severity with its impact on QOL; and (iv) the

contribution of symptom severity, anxiety and depression, and physical health status to FIQOL.

## MATERIALS and METHODS

### Study Design

In this double-blind, parallel group study which was conducted between January 2009 and April 2012, 44 women (age  $58 \pm 2$  years [Mean $\pm$ SEM]; BMI  $30.3 \pm 1.1$  kg/m<sup>2</sup>) with urge-predominant FI reported their symptoms for 4 weeks before and after they were randomized to clonidine (0.1 mg tablet, twice daily) or matched placebo. The eligibility criteria, methods, and results are detailed elsewhere (17). Enrollment was limited to patients who reported urge or combined FI on a validated questionnaire at the screening visit (18). Thereafter, only patients who had at least 4 incontinence episodes during a 4-week baseline period proceeded with the treatment phase. The Mayo Clinic Institutional Review Board approved the study, and all participants signed informed consent. It was registered at [clinicaltrials.gov](http://clinicaltrials.gov) (NCT00884832).

### Assessment of Symptoms

Patients recorded their bowel habits, including FI, on daily bowel diaries and weekly instruments. The bowel diary recorded details of every bowel movement, including the stool form (Bristol scale), presence of urgency, and characteristics of FI (17). The number of episodes of FI, number of days with FI, type and amount of leakage, and urgency were summarized per day, then per week, and finally over 8 weeks.

Every week, patients answered 5 questions that evaluated the severity of FI using the Fecal Incontinence and Constipation Assessment (FICA) instrument (1, 6). While this *questionnaire* comprehensively evaluates bowel symptoms, including constipation and FI, the letter C in the term FICA severity score is potentially misleading since the score only assesses the severity of FI, not constipation. Hence, the term FICA severity score used previously was replaced with the term Fecal Incontinence Severity Score (FISS) in this study. In addition, all patients rated their satisfaction with treatment on a 100 mm visual analog scale (VAS), from 'not satisfied at all' (on the left) to 'completely satisfied' (on the right). The FISI and QOL scales were filled at weeks 4 and 8 (19, 20). The Hospital Anxiety and Depression (HAD) and the 36-Item Short Form Health Survey (SF-36) questionnaires were assessed at baseline (21, 22).

### Statistical Analysis

Univariate associations between FI parameters were assessed with Spearman correlation coefficients and the Kruskal Wallis test. The relationship between treatment-related changes in the number of episodes and proportion of days expressed in quartiles (ordinal scale) versus corresponding changes in the FISS and VAS symptom scores was evaluated with an analysis of covariance incorporating specific contrasts for trends. Univariate associations between various variables, e.g., FI symptom severity versus QOL, were assessed with Spearman correlation coefficients. Multiple variable models assessed whether the inter-patient variation in FISS and VAS relief scores could be explained with daily bowel diaries

and, separately, the extent to which selected variables (e.g., FISS FI symptom severity score and HAD depression score) explained the FIQOL score.

To ascertain if the observed changes in the daily diary parameters (i.e., number of episodes or proportion of days with FI) were clinically important, we evaluated whether these changes were associated with a clinically-important improvement in the FISS and VAS relief scores. For each measure, the MCID was estimated with 2 accepted thresholds: 0.5 times the baseline SD (standard deviation) and the standard error of measurement, i.e., baseline SD multiplied by the square root of  $(1 - \text{Cronbach } \alpha)$  (12, 23). The relationship between the change (after – before) in the frequency of FI and the proportion of patients who reported an improvement greater than the MCID in the FISS and VAS relief scores was assessed.

## RESULTS

### Study Participants

At baseline, the FISS indicated that FI symptoms were mild (n=4), moderate (n=20), or severe (n=20). Twenty-nine patients had isolated-urge FI, and 15 had features of urge and passive FI. The overall FISS during the first 4 weeks (no treatment) and last 4 weeks (placebo or clonidine) was  $8.7 \pm 0.3$  and  $7.2 \pm 0.4$ , respectively.

### Comparison of FI Endpoints Assessed with Daily Diaries and Weekly Questionnaires over 8 weeks (i.e., before and during treatment)

The number of FI episodes, volume of FI, and severity of urgency evaluated with weekly questionnaires was correlated with the corresponding endpoint assessed with daily diaries (Figure 1). The frequency of FI assessed with weekly questionnaires was also correlated with the proportion of days with FI evaluated with a daily diary ( $r=0.91$ ,  $P<.0001$ ).

Assessments of the *changes* in symptom severity measured with daily and weekly instruments expressed as continuous (Table 1) or interval variables (Figure 2) were also significantly correlated. For example, the change (pre – post) in the FISS symptom severity score was correlated ( $r=0.57$ ,  $P .001$ ) with the change (post – pre) in the VAS relief score (Table 1). On the FISS scale, a greater score denotes more severe symptoms. On the VAS relief scale, a greater score indicates greater satisfaction with symptom relief. In addition, the change in the number of episodes of FI was correlated ( $r=0.87$ ,  $P<.001$ ) with the change in the proportion of days with FI (not shown in Table 1).

There are 3 key observations in Figure 2. First, the reduction in the frequency of FI evaluated with daily diaries was significantly associated with the change in symptom severity (i.e., FISS and VAS relief scores) evaluated with weekly questionnaires. Second, the change in VAS relief scores corresponding to all 4 categories (0-24%; 25-49%; 50-74%; and >75%) in the number of episodes of FI was numerically very similar to the VAS scores for corresponding categories in the proportion of days with FI. However, for the FISS, these values were numerically different. For example, the mean change in FISS scores for a 50%-74% reduction in the proportion of days with FI and number of episodes of FI was 1.94 and 2.62, respectively.

### Are Changes in the Frequency of FI Clinically Important?

Since the SD of the FISS symptom severity score at baseline was 1.66, the MCID derived from  $0.5 \times \text{SD}$  was 0.83. The standard error of measurement was 1.01. For the VAS relief scale,  $0.5 \times \text{SD}$  was 11.6, and the standard error of measurement was 21.6.

The reduction in the frequency of FI was associated with the proportion of patients in whom the improvement exceeded the MCID rated with the FISS scale, and to a lesser extent, with the VAS relief scale (Table 2). Among patients in whom the frequency of FI declined by 50%-74% and 75%, the FISS result exceeded the MCID threshold of  $0.5 \times \text{SD}$  (i.e., in this instance a change in FISS of 0.83) in 75% and 83% of patients, respectively. With the more stringent MCID threshold of 1.01 units based on the standard error of measurement, 75% and 50% of patients, respectively, reported clinically-important improvement with 75% and 50%-74% reduction in frequency. Using the VAS relief score, associations between the FI frequency endpoints and MCID defined by  $0.5 \times \text{SD}$ , but not the standard error of measurement, were significant.

### Multiple Variable Models to Predict FI Symptom Severity and Relief Scores from Daily Diaries

In the multiple variable model, bowel symptoms recorded by daily diaries explained 71% of the inter-patient variation in the FISS results, assessed with weekly diaries (Table 3). In this model, the symptom of urgency, the weighted amount of FI, the mean number of FI episodes/day, and the mean consistency of incontinent bowel movements explained 44%, 11%, 11%, and 10% of the variance in FISS tallies, respectively. By comparison, bowel characteristics explained a smaller proportion (29%) of the inter-patient variation in the weekly VAS relief score. Here, too, the proportion of bowel movements preceded by rectal urgency was the most useful predictor and explained 11% of the variation in relief scores; the coefficient was negative because patients with more urgency reported less relief.

### Relationship Between FI Symptom Severity and QOL Scores

The severity of FI evaluated with FISS and FISI weekly questionnaires were correlated during treatment ( $r=0.38$ ,  $P=.01$ ). However, baseline scores ( $r = 0.053$ ,  $P=.76$ ) and the change from baseline ( $r = 0.11$ ,  $P=.50$ ) were not correlated.

At baseline, correlations between FISI and the 4 subscales of FIQOL ([1] lifestyle; [2] coping/behavior; [3] depression/self-perception; and [4] embarrassment) were not significant (data not shown). After treatment, correlations between FISI and FIQOL were significant for subscales 1 ( $r = -0.44$ ,  $P=.003$ ) and 4 ( $r = -0.31$ ,  $P=.045$ ) but not 2 and 3.

At baseline, correlations between FISS and the 4 subscales of FIQOL were not significant (data not shown). After treatment, correlations between FISS and subscales 1 ( $r = -0.35$ ,  $P=.02$ ) and 2 ( $r = -0.44$ ,  $P=.004$ ), but not 3 or 4, were significant. The correlations between treatment-related changes (after - before treatment) in FIQOL subscale 4 and FISS, and separately with FISI, were borderline significant (Table 4). However, correlations between changes in FISI and FISS and the other 3 subscales were not significant.

## Relationship Between Mental and Physical Health Function Scores, FI Symptom Severity and QOL

Except for the coping behavior scale after treatment, the scores for depression, computed by the HAD instrument, were inversely correlated with all 4 subscales of the FIQOL before and after treatment (Table 5). These inverse correlations imply that greater scores (i.e., more severe) for depression were associated with poorer FIQOL.

Scores for several of the 8 scales for mental and physical health of the SF-36 instrument were significantly correlated with FIQOL scores (Table 5). In particular, baseline SF36 scores for general health, physical role functioning, and social functioning were directly correlated with virtually all 4 subscores of the FIQOL instrument at baseline and during therapy, implying that better scores on the SF-36 QOL instrument were associated with better FIQOL. In comparison, the other SF-36 scales were either weakly or not correlated with the FIQOL. Likewise, the SF-36 scores were not correlated with the change in FIQOL during therapy. Also, correlations between both the SF-36 scales and FI symptom severity evaluated with the FISS symptom severity scale and the FISI instrument at baseline and after therapy were not significant (data not shown).

In the multiple linear regression models, FI symptoms and other variables explained between 18% and 52% of the inter-subject variation in the FIQOL scores (Table 6). Of note, with one exception (the FIQOL subscore for coping/behavior after therapy), the FISS did not explain the variation in FIQOL scores. Conversely, the SF-36 scores for physical and mental health were the strongest predictors of this variation. Similar results were obtained when the FISS score was substituted for the FISI score in these models (data not shown).

## DISCUSSION

Daily and weekly instruments are used to characterize the endpoints of FI: frequency, amount, consistency, and circumstances surrounding FI. There are 3 main observations in this study. First, several measures of symptom severity evaluated with daily and weekly instruments were strongly correlated; in particular, the correlation coefficients of daily versus weekly measures of FI frequency were 0.9. The FISS scale derived from the weekly questionnaire was correlated with the individual FI parameters in the daily diary and with the VAS relief score. Taken together, these observations confirm the construct validity of both instruments and demonstrate that weekly instruments are reliable for evaluating the severity of FI. Second, 75% of patients with a 50% or greater reduction in the number of FI episodes experienced a clinically-significant improvement, which validates this endpoint in therapeutic trials. Third, in addition to symptoms, physical and mental health independently explain FIQOL.

Currently, a 50% reduction in the number of episodes or days with FI is the primary outcome measure in therapeutic trials of FI (24). The Food and Drug Administration deems a 50% reduction to be a “clinically meaningful” improvement in urinary incontinence (13). A similar threshold has also been used to identify clinically significant responders in irritable bowel syndrome (25). However, it has been suggested that a 50% reduction in the frequency of FI may not be clinically meaningful (11). Addressing this question, we



observed significant associations between changes in the frequency of FI, expressed on an ordinal scale, and (i) corresponding changes in the global endpoints assessed with weekly diaries (i.e., the FISS symptom severity score and VAS relief score) and (ii) the proportion of patients in whom the FISS symptom severity score and the VAS relief score exceeded the MCID. We used distribution-based methods to assess the MCID in FI because there are no established anchor-based approaches to determine the same. There is growing consensus that a change of  $0.5 \times \text{SD}$  is a conservative estimate that is likely to be clinically important across different patient-reported questionnaires (26, 27). Indeed, it has been proposed that, in the absence of other information,  $0.5 \times \text{SD}$  is a reasonable estimate of meaningful effect (27). Among patients in whom the frequency of FI declined by 50%-74% and 75%, the FISS result exceeded the MCID in 75% and 83% of patients, respectively. The 50% threshold used in clinical trials includes patients with 50%-74% and 75% improvement. While a 75% reduction in the frequency of FI is intuitively preferable to reduction of 50%-74%, these observations support the currently-used outcome measures that are based on a 50% reduction in the frequency of FI. Because the numbers are small, we recognize that further studies are necessary to determine whether the proportion of patients who have clinically-significant improvement differs between the 50%-74% and 75% categories.

The daily diary endpoints explained 71% of the variation in FI symptom severity assessed with weekly questionnaires. The proportion of bowel movements that were associated with urgency was the strongest predictor of the inter-subject variation in weekly FISS and VAS symptom relief scores. This observation underscores the contribution of rectal urgency to the severity of FI. Only 2 of the several FI symptom severity instruments (the St Marks scale and the FISS) incorporate the symptom of rectal urgency, and only the FISS scale includes the volume of leakage (28). It is conceivable that patients completed their weekly questionnaires while reviewing their daily diaries. To the extent this occurred, this would exaggerate the agreement between daily and weekly instruments.

These observations confirm the validity of weekly diaries for evaluating the severity of FI. Weekly diaries are particularly useful for evaluating long-term effectiveness, which is necessary to adequately assess efficacy. However, weekly instruments are potentially susceptible to recall bias and perhaps do not provide a refined assessment of the relationship between stool consistency and FI (9). Perhaps one approach to balancing the need to accurately characterize symptoms while minimizing the burden for patients is to gather data every week for 3 months and then administer questionnaires for 4 weeks at periodic intervals, e.g., every 6 months for up to 5 years.

The changes in symptom severity measured with the FISI and FISS were not significantly correlated with the corresponding changes in the QOL in this 4-week study. However, there was a trend for an association between the change in the embarrassment subscore of the FIQOL scale and the improvement in the FISS and FISI scores. Since many facets of QOL involve the ability to engage in daily activities, it is conceivable that more time is required before women regain the confidence to engage in activities, and hence improve their lifestyle. However, even at 1, 2, 3, and 4 years after SNS, only approximately 5% to 15% of the variance in the changes on the FIQOL scales were explained by changes in the symptom severity of FI (14). In the dextranomer study, the improvement in symptom severity was not

accompanied by an improvement in QOL (29). Consistent with the biopsychosocial model, it is conceivable that other factors (e.g., depression), which have been explored in other conditions, such as functional dyspepsia, may explain the limited correlation between symptom severity and QOL in FI (30). Indeed, the HAD depression score and SF-36 scores for physical and mental health at baseline explained a greater proportion of the inter-subject differences in pre- and post-treatment FIQOL scores than the FISS FI symptom severity score. Hence, future therapeutic trials in FI should also assess psychosocial issues and overall health before and after therapy.

Among 44 subjects, 27% had a normal BMI, 30% were overweight, 25% were obese, and 18% were severely obese. Obesity and bariatric surgery are risk factors for FI (31-33). Obesity is associated with faster colonic transit and diarrhea as also with greater intra-abdominal pressure, all of which may predispose to FI (32).

These observations need to be confirmed by future studies. A larger study in which subjects are exposed to an efficacious intervention would increase the number of patients who experienced improved symptoms and QOL in each stratum (e.g., 50%-74%). The effect of response bias on the comparison between daily and weekly questionnaires can be minimized by obscuring responses to daily diaries once they have been completed. In summary, a majority of women who report a 50% reduction in FI frequency experience a clinically important improvement. Weekly questionnaires accurately assess the severity of FI. Self-reported physical and mental health explained a greater proportion of the variance in FIQOL than FI symptom severity.

## ACKNOWLEDGMENTS

**Funding:** This study was supported by US Public Health Service National Institutes of Health grant R01 DK78924. This project was supported by grant number UL1 TR000135 from the National Center for Advancing Translational Sciences. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institutes of Health.

## Abbreviations

<b>FI</b>	fecal incontinence
<b>FICA</b>	Fecal Incontinence and Constipation Assessment
<b>FISS</b>	Fecal Incontinence Severity Score
<b>FISI</b>	Fecal Incontinence Severity Index
<b>HAD</b>	Hospital Anxiety and Depression
<b>MCID</b>	minimum clinically important difference
<b>QOL</b>	quality of life
<b>SNS</b>	sacral nerve stimulation
<b>VAS</b>	visual analog scale



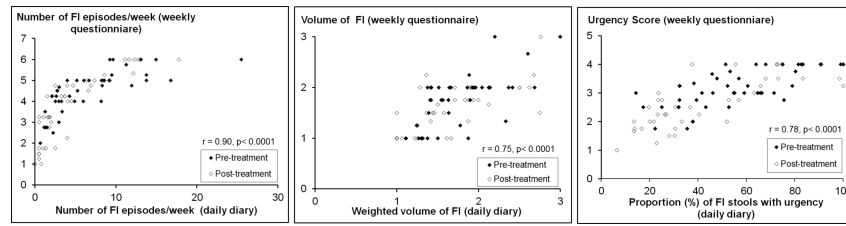
## REFERENCES

1. Bharucha AE, Zinsmeister AR, Locke GR, Seide B, McKeon K, Schleck CD, Melton LJI. Prevalence and burden of fecal incontinence: A population based study in women. *Gastroenterology*. 2005; 129:42–9. [PubMed: 16012933]
2. Quander CR, Morris MC, Melson J, Bienias JL, Evans DA. Prevalence of and factors associated with fecal incontinence in a large community study of older individuals. *Am J Gastroenterol*. 2005; 100:905–9. [PubMed: 15784039]
3. Perry S, Shaw C, McGrother C, Matthews RJ, Assassa RP, Dallosso H, Williams K, Brittain KR, et al. Prevalence of faecal incontinence in adults aged 40 years or more living in the community. *Gut*. 2002; 50:480–4. [PubMed: 11889066]
4. Varma MG, Brown JS, Creasman JM, Thom DH, Van Den Eeden SK, Beattie MS, Subak LL. Reproductive Risks for Incontinence Study at Kaiser Research G. Fecal incontinence in females older than aged 40 years: who is at risk? *Dis Colon Rectum*. 2006; 49:841–51. [PubMed: 16741640]
5. Whitehead WE, Borrud L, Goode PS, Meikle S, Mueller ER, Tuteja A, Weidner A, Weinstein M, et al. Fecal incontinence in US adults: epidemiology and risk factors. *Gastroenterology*. 2009; 137:512–7. [PubMed: 19410574]
6. Bharucha AE, Zinsmeister AR, Locke GR, Schleck C, McKeon K, Melton LJ. Symptoms and quality of life in community women with fecal incontinence. *Clin Gastroenterol Hepatol*. 2006; 4:1004–9. [PubMed: 16630766]
7. Omar MI, Alexander CE. Drug treatment for faecal incontinence in adults. *Cochrane Database Syst Rev*. 2013; 6:CD002116. [PubMed: 23757096]
8. Hull T, Giese C, Wexner SD, Mellgren A, Devroede G, Madoff RD, Stromberg K, Collier JA, et al. Long-term durability of sacral nerve stimulation therapy for chronic fecal incontinence. *Dis Colon Rectum*. 2013; 56:234–45. [PubMed: 23303153]
9. Bharucha AE, Seide B, Zinsmeister AR, Melton JL. Relation of bowel habits to fecal incontinence in women. *Am J Gastroenterol*. 2008; 103:1470–5. [PubMed: 18510612]
10. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials*. 2003; 24:182–99. [PubMed: 12689739]
11. Wald A. Clonidine and botulinum toxin: a tale of two treatments. *Clin Gastroenterol Hepatol*. 2014; 12:852–3. [PubMed: 23973926]
12. 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]
13. Barber MD, Spino C, Janz NK, Brubaker L, Nygaard I, Nager CW, Wheeler TL, Pelvic Floor Disorders N. The minimum important differences for the urinary scales of the Pelvic Floor Distress Inventory and Pelvic Floor Impact Questionnaire. *Am J Obstet Gynecol*. 2009; 200:580.e1–7. [PubMed: 19375574]
14. Devroede G, Giese C, Wexner SD, Mellgren A, Collier JA, Madoff RD, Hull T, Stromberg K, et al. Quality of life is markedly improved in patients with fecal incontinence after sacral nerve stimulation. *Female Pelvic Med Reconstr Surg*. 2012; 18:103–12. [PubMed: 22453321]
15. Koloski NA, Jones M, Kalantar J, Weltman M, Zaguirre J, Talley NJ. Psychological impact and risk factors associated with new onset fecal incontinence. *J Psychosom Res*. 2012; 73:464–8. [PubMed: 23148816]
16. Smith TM, Menees SB, Xu X, Saad RJ, Chey WD, Fenner DE. Factors associated with quality of life among women with fecal incontinence. *Int Urogynecol J*. 2013; 24:493–9. [PubMed: 22806489]
17. Bharucha AE, Fletcher JG, Camilleri M, Edge J, Carlson P, Zinsmeister AR. Effects of clonidine in women with fecal incontinence. *Clin Gastroenterol Hepatol*. 2014; 12:843–51 e2. quiz e44. [PubMed: 23891925]
18. Bharucha AE, Locke GR, Seide B, Zinsmeister AR. A New Questionnaire for Constipation and Fecal Incontinence. *Aliment Pharmacol Ther*. 2004; 20:355–64. [PubMed: 15274673]
19. Rockwood TH, Church JM, Fleshman JW, Kane RL, Mavrantonis C, Thorson AG, Wexner SD, Bliss D, et al. Fecal incontinence quality of life scale: quality of life instrument for patients with fecal incontinence. *Dis Colon Rectum*. 2000; 43:9–16. discussion -7. [PubMed: 10813117]

20. Rockwood TH, Church JM, Fleshman JW, Kane RL, Mavrantonis C, Thorson AG, Wexner SD, Bliss D, et al. Patient and surgeon ranking of the severity of symptoms associated with fecal incontinence: the fecal incontinence severity index. *Dis Colon Rectum*. 1999; 42:1525–32. [PubMed: 10613469]
21. Brazier JE, Harper R, Jones NM, O’Cathain A, Thomas KJ, Usherwood T, Westlake L. Validating the SF-36 health survey questionnaire: new outcome measure for primary care. *BMJ*. 1992; 305:160–4. [PubMed: 1285753]
22. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983; 67:361–70. [PubMed: 6880820]
23. 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]
24. Wexner SD, Collier JA, Devroede G, Hull T, McCallum R, Chan M, Ayscue JM, Shobeiri AS, et al. Sacral nerve stimulation for fecal incontinence: results of a 120-patient prospective multicenter study. *Ann Surg*. 2010; 251:441–9. [PubMed: 20160636]
25. Whitehead WE, Palsson OS, Levy RL, Feld AD, VonKorff M, Turner M. Reports of “satisfactory relief” by IBS patients receiving usual medical care are confounded by baseline symptom severity and do not accurately reflect symptom improvement. *Am J Gastroenterol*. 2006; 101:1057–65. [PubMed: 16573774]
26. 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]
27. 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]
28. Vaizey CJ, Carapeti E, Cahill JA, Kamm MA. Prospective comparison of faecal incontinence grading systems. *Gut*. 1999; 44:77–80. [PubMed: 9862829]
29. Graf W, Mellgren A, Matzel KE, Hull T, Johansson C, Bernstein M. Efficacy of dextranomer in stabilised hyaluronic acid for treatment of faecal incontinence: a randomised, sham-controlled trial. *Lancet*. 2011; 377:997–1003. [PubMed: 21420555]
30. Bharucha AE, Camilleri M, Burton DD, Thieke SL, Feuerhak KJ, Basu A, Zinsmeister AR. Increased nutrient sensitivity and plasma concentrations of enteric hormones during duodenal nutrient infusion in functional dyspepsia. *Am J Gastroenterol*. 2014; 109:1910–20. [PubMed: 25403365]
31. Bharucha AE, Fletcher JG, Melton LJ 3rd, Zinsmeister AR. Obstetric Trauma, Pelvic Floor Injury And Fecal Incontinence: A Population-Based Case-Control Study. *Am J Gastroenterol*. 2012; 107:902–11. [PubMed: 22415196]
32. Bharucha AE. Incontinence - An Underappreciated Problem In Obesity And Bariatric Surgery. *Dig Dis Sci*. 2010; 55:2428–30. [PubMed: 20521110]
33. Roberson EN, Gould JC, Wald A. Urinary and Fecal Incontinence after Bariatric Surgery. *Dig Dis Sci*. 2010; 55:2606–13. [PubMed: 20393881]

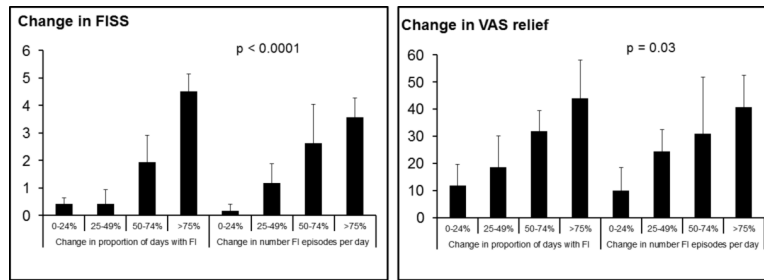
**KEY MESSAGES**

- This study suggests that a greater than 50% reduction in the frequency of FI is clinically significant, that symptom severity evaluated with daily and weekly instruments are strongly correlated, and that depression as well as physical and mental health independently explain QOL in FI.
- The aims of this study were to further validate the endpoints used to assess symptom severity and QOL in clinical trials of FI
- FI severity was assessed with daily bowel diaries and periodic questionnaires (Fecal Incontinence Severity score [FISS], FIQOL, 36-Item Short Form Health Survey [SF-36], and Hospital Anxiety and Depression scales) for 4 weeks before and during double-blind randomization to placebo or clonidine in 44 women with FI.
- These findings are important because they substantiate the endpoints used to evaluate symptom severity in FI, suggest that weekly instruments may be useful to assess symptom severity, and highlight the need to incorporate psychological variables when interpreting the effect of FI on QOL.



**Figure 1. Correlations Between Fecal Incontinence (FI) Characteristics Assessed with Daily Diaries and Weekly Questionnaires**

For the volume of leakage (center panel), scores represent (1) staining only; (2) a moderate amount of leakage, i.e., more than a stain but less than a complete bowel movement; and (3) gross leakage requiring change of garments. The legend for the urgency score (weekly questionnaire, right panel) are: (1) neither urge nor passive FI; (2) passive FI only; (3) urge FI; and (4) urge and passive FI.



**Figure 2. Comparison of Treatment-Related Endpoints Assessed with Daily Diaries (Interval Scale) and Weekly Questionnaires**

Observed significant correlations between daily endpoint (change, after – before, in number FI episodes) and weekly endpoint (change in FISS and VAS relief score).

**Table 1**

Comparison of Different Measures of Fecal Incontinence Assessed with Daily Diaries (Continuous Scale) and Weekly Questionnaires <sup>a</sup>

Parameter	Change in number of episodes with FI (Daily Diary)	Change in proportion of days with FI (Daily Diary)	Change in FISS <sup>b</sup>	Change in FI VAS relief score <sup>b,f</sup>
Change in FI frequency <sup>b</sup>	0.42 <sup>c</sup>	0.68 <sup>c</sup>	0.82 <sup>c</sup>	0.58 <sup>c</sup>
Change in FISS <sup>b</sup>	0.36 <sup>d</sup>	0.41 <sup>e</sup>		0.57 <sup>c</sup>
Change in VAS relief score <sup>b,f</sup>	0.46 <sup>e</sup>	0.46 <sup>e</sup>	0.57 <sup>c</sup>	

Abbreviations: FI, fecal incontinence; FISS, Fecal Incontinence Severity Score; VAS, visual analog scale

<sup>a</sup>Spearman correlation coefficients (*P* values). All changes are Pre [4 week average] – Post [4 week average] except where indicated otherwise

<sup>b</sup>Derived from weekly questionnaires

<sup>c</sup>*P* .001

<sup>d</sup>*P* < .05

<sup>e</sup>*P* .01

<sup>f</sup>Post – pre.



**Table 2**  
 Comparison of Changes in Frequency of FI on Bowel Diary versus Minimally Important Differences in FISS and VAS Relief Scores

	Number (row %) of subjects with improvement in FISS > MCID		Number (row %) of subjects with improvement in VAS relief score > MCID	
	N	>0.83 <sup>a</sup>	N	>11.6 <sup>a</sup>
Change in proportion of days with FI	0-24%	6 (40%)	14	6 (43%)
	25-49%	4 (40%)	10	5 (50%)
	50-74%	4 (67%)	6	6 (100%)
	75%	9 (90%)	10	8 (80%)
<i>P</i> value <sup>c</sup>		.050		.046
Change in number FI episodes per day	0-24%	5 (31%)	15	5 (33%)
	25-49%	5 (56%)	9	7 (78%)
	50-74%	3 (75%)	4	3 (75%)
	75%	10 (83%)	12	10 (83%)
<i>P</i> value <sup>c</sup>		.037		.033

Abbreviations: FI, fecal incontinence; FISS, Fecal Incontinence Severity Score; MCID, Minimal Important Differences; VAS, visual analog scale

<sup>a</sup>0.5SD

<sup>b</sup>Standard error of measurement.

<sup>c</sup>*P*value for association with group status

**Table 3**  
Multiple Variable Predictor Models for Predicting Weekly Assessments from Daily Bowel Diary Parameters

Daily diary parameters	Weekly FISS (8 weeks)		Weekly VAS Relief Score (between weeks 4 and 8)	
	R <sup>2</sup>	Coefficient (SE)	R <sup>2</sup>	Coefficient (SE)
Proportion of days with FI	0.04	1.87 (1.60)	0.01	-26.30 (42.71)
Mean number of FI episodes/day	0.11 <sup>a</sup>	1.52 (0.71)	0.003	-7.17 (20.89)
Mean consistency of incontinent bowel movements	0.10 <sup>a</sup>	0.31 (0.15)	0.05	3.82 (2.95)
Weighted amount of FI <sup>c</sup>	0.11 <sup>a</sup>	4.34 (2.05)	0.002	-11.80 (42.46)
Proportion of incontinent bowel movements with urgency	0.44 <sup>b</sup>	3.96 (0.74)	0.11 <sup>a</sup>	-29.76 (14.06)
Variance explained (R <sup>2</sup> )	71%		29%	

Abbreviations: FI, fecal incontinence; FISS, Fecal Incontinence Severity Score; VAS, visual analog scale

<sup>a</sup>  $P < .05$

<sup>b</sup>  $P < .0001$

<sup>c</sup> Weighted amount of FI =  $(1 * \text{proportion of staining episodes} + 2 * \text{proportion of soiling episodes} + 3 * \text{proportion of episodes requiring change of clothes}) / 6$ ;

**Table 4**Relationship Between Change in FISS and FISI Symptom Severity and Quality of Life Scores <sup>a</sup>

Change in FISI QOL subscores	Change in FISS (Post-Pre)	Change in FISI (Post-Pre)
Lifestyle (subscale 1)	-0.094 (0.58)	-0.18 (0.28)
Coping/behavior (subscale 2)	-0.28 (0.10)	-0.21 (0.21)
Depression/Self Perception (subscale 3)	-0.12 (0.48)	-0.040 (0.81)
Embarrassment (subscale 4)	-0.31 (0.06)	-0.32 (0.06)

Abbreviations: FISI, Fecal Incontinence Severity Index; FISS, Fecal Incontinence Severity Score; QOL, Quality of Life

Values are Spearman correlation coefficients (*P* values)<sup>a</sup>Changes are (After – before)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 5**

Univariate Analysis of the Relationship Between Mental and Physical Health Function Scores and FIQOL

Parameter	FIQOL scale <sup>a</sup>							
	Lifestyle (subscale 1)		Coping/behavior (subscale 2)		Depression/self Perception (subscale 3)		Embarrassment (subscale 4)	
	Baseline	After	Baseline	After	Baseline	After	Baseline	After
HAD Anxiety	-0.15	-0.056	-0.15	0.057	-0.39 <sup>b</sup>	-0.17	-0.20	-0.28
HAD Depression	-0.57 <sup>d</sup>	-0.41 <sup>b</sup>	-0.45 <sup>c</sup>	-0.26	-0.52 <sup>d</sup>	-0.43 <sup>c</sup>	-0.42 <sup>b</sup>	-0.34 <sup>b</sup>
Physical_Health	0.27	0.25	0.41 <sup>b</sup>	0.20	0.54 <sup>d</sup>	0.40 <sup>b</sup>	0.17	0.22
Mental_Health	0.084	-0.14	0.18	-0.18	0.39 <sup>b</sup>	0.24	0.083	-0.048

Abbreviations: FI, fecal incontinence; FIS, Fecal Incontinence Severity Index; HAD, Hospital Anxiety and Depression; QOL, quality of life

<sup>c</sup>  $P < .0001$ <sup>a</sup> Values are Spearman correlation coefficients<sup>b</sup>  $P < .05$ <sup>c</sup>  $P < .01$ <sup>d</sup>  $P < .001$

**Table 6**

Multiple Variable Analysis of Risk Factors Associated with QOL in FI <sup>a</sup>

Parameter <sup>b</sup>	FIQOL subscores									
	Lifestyle (subscale 1)		Coping/behavior (subscale 2)		Depression/Self Perception (subscale 3)		Embarrassment (subscale 4)		Total variance explained (R <sup>2</sup> %)	
	Baseline	After	Baseline	After	Baseline	After	Baseline	After		
FISS	0.013	0.012	0.0053	0.21 <sup>d</sup>	0.0032	0.0031	0.0080	0.18		
HAD Depression	0.13 <sup>c</sup>	0.018	0.014	0.003	0.013	0.061	0.034	0.028		
SF36 - Physical Health	0.11 <sup>c</sup>	0.20 <sup>d</sup>	0.20 <sup>d</sup>	0.002	0.28 <sup>d</sup>	0.036	0.11	0.020		
SF36 - Mental Health	0.022	0.14 <sup>c</sup>	0.15 <sup>c</sup>	0.052	0.40 <sup>e</sup>	0.11	0.12 <sup>c</sup>	0.068		
	31%	31%	30%	18%	52%	22%	25%	24%		

Abbreviations: HAD, Hospital Anxiety and Depression; FI, fecal incontinence; FISS, Fecal Incontinence Severity Score; FISI, Fecal Incontinence Severity Index; QOL, quality of life; SF36, 36-Item Short Form Health Survey

<sup>a</sup>Values are squared partial correlation coefficients (scale 0 to 1)

<sup>b</sup>Baseline and treatment FISS results were used for baseline and treatment models, respectively. For other parameters, the baseline scores were used.

<sup>c</sup>P < .05

<sup>d</sup>P < .01

<sup>e</sup>P < .0001