Citation: Clinical and Translational Gastroenterology (2015) 6, e70; doi:10.1038/ctg.2014.20

www.nature.com/ctg

Validity of the ≥50% Response Threshold in Treatment With NASHA/Dx Injection Therapy for Fecal Incontinence

Jaime E. Sanchez, MD, MSPH[1], Darren M. Brenner, MD[2], Howard Franklin, MD[3], Jing Yu, PhD[3], Andrew C. Barrett, PhD[3] and Craig Paterson, MD[3]

OBJECTIVES: Many fecal incontinence (FI) studies define primary efficacy outcome as a decrease from baseline of ≥50% in the number of FI episodes; this threshold has never been validated. We aimed to establish the validity and responsiveness of ≥50% reduction in FI episodes (responder$_{50}$) as the threshold indicative of clinically meaningful response.

METHODS: Adults with a Cleveland Clinic Florida fecal incontinence score ≥10 were randomized to receive nonanimal stabilized hyaluronic acid/dextranomer (NASHA/Dx) injection or sham treatment in a 6-month trial. Validity and responsiveness of the primary end point were evaluated *post hoc*. The data were compared using different thresholds for defining a responder for a number of end points.

RESULTS: Data from 206 patients (NASHA/Dx, n = 136; sham, n = 70) were evaluated. Incremental patient response threshold increases showed that although the percentage of patients who achieved response decreased with increasing threshold, the difference between treatments remained significant up to an 80% response threshold (NASHA/Dx, 23%; sham, 10%; $P = 0.02$). Response thresholds between 40% and 80% demonstrated evidence for convergent validity, with the strongest correlation with the number of FI episodes, the number of FI episodes when the patient was awake, and the number of FI-free days observed at ≥40% and ≥50% thresholds. Further examination of the responder$_{50}$ threshold indicated that, regardless of treatment (NASHA/Dx or sham), responders performed significantly better than nonresponders on nearly all secondary efficacy end points.

CONCLUSION: This study demonstrates the responsiveness, validity, and clinical applicability of the ≥50% response threshold in clinical studies of patients with FI receiving treatment with NASHA/Dx.

Clinical and Translational Gastroenterology (2015) **6,** e70; doi:10.1038/ctg.2014.20; published online 15 January 2015
Subject Category: Functional GI Disorders

INTRODUCTION

Fecal incontinence (FI) is a multifactorial disorder, affecting approximately 8.4% of noninstitutionalized adults in the United States.[1–3] Patients with FI are not only affected by the physical symptoms (e.g., fecal seepage, perianal dermatitis) but also by the psychosocial aspects of the condition (e.g., embarrassment, restriction of activity).[4–6] Clinical symptoms of FI are variable, with some patients experiencing as few as 1–3 FI episodes per month and others experiencing ≥1 FI episode per day.[3] Patients with FI experience the inability to control bowel emptying and may report their stool consistency to be watery, solid, or a combination of both.[3,6]

Patients with FI have a number of treatment options available, including dietary modification, pharmacological therapies, physical therapy/biofeedback, injectable bulking agents, and invasive alternatives, such as sphincteroplasty or sacral nerve stimulation.[7,8] Clinical trials of these therapies for patients with FI have been inconsistent in defining response to treatment.[9] Even when investigators define response as a decrease from baseline in the number of FI episodes, there is variability in the threshold for response. For example, a clinical study of biofeedback for treatment of FI defined treatment success as a ≥80% overall decrease in the number of FI episodes.[10] By contrast, clinical trials of injectable bulking agents in patients with FI often have defined treatment response as a decrease from baseline of ≥50% in the number of FI episodes at a given time point.[11–14] Determination of a ≥50% threshold for response has typically been based on previous data from patients with FI and on the assumption that this cutoff is clinically meaningful for patients.[11,12] The goal of this study was to establish the validity and responsiveness of the ≥50% response threshold for assessing the efficacy of treatment with nonanimal stabilized hyaluronic acid/dextranomer (NASHA/Dx) in patients with FI.

METHODS

Patients and study design. Details of the study population, inclusion and exclusion criteria, and study design have been described previously.[14] Patients aged 18–75 years with FI

[1]Division of Colon and Rectal Surgery, University of South Florida Morsani College of Medicine, Tampa, Florida, USA; [2]Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA and [3]Salix Pharmaceuticals Inc., Raleigh, North Carolina, USA
Correspondence: Jaime E. Sanchez, MD, MSPH, University of South Florida Morsani College of Medicine, Division of Colon and Rectal Surgery, 1 Tampa General Circle, Suite F-145, Tampa, Florida 33606, USA. E-mail: jasanche@health.usf.edu
Received 5 August 2014; accepted 17 November 2014

(measured as Cleveland Clinic Florida fecal incontinence score (CCFIS) ≥ 10 and at least 4 FI episodes over a 2-week period) from the United States and Europe were randomized (2:1) to receive anal injections of NASHA/Dx or sham treatment, respectively. Immediately prior to receiving therapy, patients received a cleansing enema. An anoscope was used to administer a 1-ml injection of NASHA/Dx into each quadrant of the submucosa, approximately 5 mm above the dentate line. Patients randomized to sham treatment received the same procedure without injection of any substance. Patients were followed for 1 year. Investigators administering the study treatment were not blinded, but patients and investigators conducting clinical assessments at each of the treatment centers were blinded to treatment for the first 6 months. The study was approved by the institutional review boards and ethics committees of all the participating facilities, and all patients provided written informed consent. This trial was registered in the ClinicalTrials.gov database (NCT00605826).

Assessments. The primary efficacy end point was response to treatment, defined as a ≥ 50% decrease in the number of FI episodes during a 2-week period at 6 months posttreatment compared with baseline.[14] Secondary efficacy end points included the change from baseline in the number of FI-free days and the number of FI episodes at 6 months and the change from baseline in CCFIS at 6 months.

Statistical analyses. Efficacy was evaluated in the intention-to-treat population, defined as randomized patients who received ≥ 1 injection.[14] Responder threshold analysis was conducted by generating point estimates and corresponding odds ratios (OR) and 95% confidence intervals (CIs) using a logistic model with covariates (i.e., baseline number of FI episodes, sex, and treatment center). The primary end point of the ≥ 50% response threshold was evaluated by validity and responsiveness analysis. The validity analysis determined whether an end point accurately reflected the concept it was intended to measure. Convergent validity and divergent validity are two typical methods of assessing validity. These types of validity are based on correlations among measures. A correlation of ≥ 0.4 was considered evidence for convergent validity, and a correlation < 0.3 was considered evidence for divergent validity. Correlations between 0.3 and 0.4 were considered as no evidence to establish or dismiss convergent or divergent validity.[15–17] Spearman's rank correlation (r_s) coefficients were calculated to assess the correlations among the primary end point and secondary efficacy end points. Responsiveness, the ability of an end point to detect small but clinically important changes, was assessed by conducting a Wilcoxon's rank-sum test to compare the mean change from baseline to 6 months in secondary efficacy end points in patients who achieved ≥ 50% reduction in the number of FI episodes with patients who did not achieve ≥ 50% reduction in FI episodes (regardless of treatment group assignment). To determine the minimal clinically important difference for the primary end point, one-half the s.d. of the baseline number of FI episodes was divided by the mean number of FI episodes (excluding outliers).[18] For this calculation, major outliers were excluded (i.e., values outside the boundaries of the outer fences, which were calculated as (first quartile – 3 × interquartile range) and (third quartile+3 × interquartile range)).

Table 1 Summary of the demographic and baseline disease characteristics[14]

Characteristic	NASHA/Dx ($n = 136$)	Sham ($n = 70$)
Age, years, mean (range)	61.8 (55.5–68.3)	60.1 (51.3–66.7)
Female, *n* (%)	122 (90)	61 (87)
Baseline FI episodes, median (range)	15 (9.6–27.5)	12.5 (8.0–28.0)
Baseline CCFIS, median (range)	14.0 (12.0–16.0)	13.0 (12.0–15.0)
Duration of symptoms, *n* (%)		
1–5 years	65 (47.8)	35 (50)
> 5 years	71 (52.2)	35 (50)

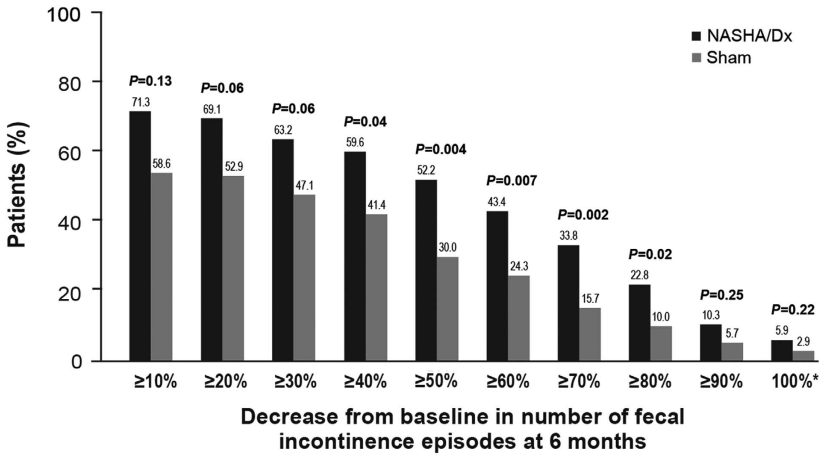
CCFIS, Cleveland Clinic Florida fecal incontinence score; FI, fecal incontinence; NASHA/Dx, nonanimal stabilized hyaluronic acid/dextranomer.
Adapted with permisson of Lancet Publishing Group from Graf *et al.*[14] Permission conveyed through Copyright Clearance Center, Inc.

RESULTS

Patient population. A total of 206 patients (NASHA/Dx, $n = 136$; sham, $n = 70$) were included in the intention-to-treat population.[14] Demographic and baseline characteristics were comparable between the two treatment groups (Table 1) and have been previously reported.[14] At baseline, each treatment group had a comparable median number of FI episodes (NASHA/Dx, $n = 15$; sham, $n = 12.5$; $P = 0.62$) and median CCFIS (NASHA/Dx, 14.0; sham, 13.0; $P = 0.19$).

Efficacy. A significantly greater percentage of patients receiving NASHA/Dx achieved the primary efficacy end point (i.e., a ≥ 50% decrease from baseline in the number of FI episodes at 6 months) compared with those receiving sham treatment (52.7% vs. 32.1%, respectively; $P = 0.009$).[14] *Post hoc* analysis of the ≥ 50% response threshold, which differed from the published result in that the treatment center effect was removed, replicated this finding (NASHA/Dx vs. sham treatment, 52.2% vs. 30.0%; $P = 0.004$). Significant differences between the treatment arms were maintained when the threshold for response was incrementally increased from 40% to 80% (Figure 1). The largest absolute difference relative to sham treatment was observed at the 50% response threshold ($\Delta = 22.2\%$). The minimal clinically important difference for the primary outcome was calculated to be 35% (see Methods section). Across response thresholds ranging from ≥ 40% to ≥ 80%, the ORs ranged from 1.95 to 3.20, with 95% CIs that did not overlap 1.0. However, with each incremental increase in threshold from 40%, there was a concomitant decrease of approximately 7–11% in the percentage of patients with response to NASHA/Dx. A similar decrease was observed for sham treatment.

To evaluate the validity of end points using different response thresholds, r_s coefficients were calculated between the response thresholds (between ≥ 40% and ≥ 80%) and secondary efficacy end points. In general, the ≥ 40% response threshold and the ≥ 50% response threshold (primary end



	≥10%	≥20%	≥30%	≥40%	≥50%	≥60%	≥70%	≥80%	≥90%	100%*
Difference relative to sham (%)	12.7	16.2	16.1	18.2	22.2	19.1	18.1	12.8	4.6	3.0
Odds ratio	1.63	1.83	1.84	1.95	2.56	2.49	3.20	3.04	1.97	2.58
95% CI	0.86, 3.10	0.97, 3.44	0.98, 3.44	1.05, 3.64	1.34, 4.90	1.25, 4.94	1.46, 7.03	1.16, 7.96	0.58, 6.71	0.50, 13.26

Figure 1 Patient response (i.e., decrease from baseline in the number of fecal incontinence episodes) to NASHA/Dx or sham treatment at 6 months at different thresholds. *In the analysis of the 100% threshold, the treatment center effect was removed from the logistic regression model to resolve the issue of model convergence. CI, confidence interval; NASHA/Dx, nonanimal stabilized hyaluronic acid/dextranomer.

Table 2 Correlation of responder thresholds with patient-reported efficacy outcome measures at 6 months

	Responder thresholds[a,b]				
Change from baseline in efficacy outcome	**≥ 40%**	**≥ 50%**	**≥ 60%**	**≥ 70%**	**≥ 80%**
CCFIS total score	0.42	0.38	0.39	0.36	0.40
Number of FI episodes	0.79	0.71	0.64	0.59	0.48
Number of FI episodes when awake	0.78	0.70	0.62	0.59	0.48
Number of FI episodes when asleep	0.25	0.25	0.30	0.23	0.18
Number of FI episodes with solid stools	0.46	0.43	0.37	0.34	0.26
Number of FI episodes with loose stools	0.53	0.44	0.40	0.36	0.36
Number of controlled bowel-emptying episodes with fecal urgency	0.07	0.15	0.18	0.12	0.11
Number of controlled bowel-emptying episodes	− 0.07	− 0.05	− 0.05	− 0.02	− 0.08
Number of FI-free days	− 0.71	− 0.72	− 0.71	− 0.64	− 0.59
FIQOL coping and behavior score	− 0.21	− 0.29	− 0.31	− 0.30	− 0.32
FIQOL lifestyle score	− 0.24	− 0.24	− 0.27	− 0.25	− 0.28
FIQOL depression and self perception score	− 0.15	− 0.18	− 0.17	− 0.24	− 0.29
FIQOL embarrassment score	− 0.20	− 0.19	− 0.23	− 0.25	− 0.26

CCFIS, Cleveland Clinic Florida fecal incontinence score; FI, fecal incontinence; FIQOL, fecal incontinence quality of life; NASHA/Dx, nonanimal stabilized hyaluronic acid/dextranomer.
[a]Minimum percentage of decrease from baseline in the number of FI episodes.
[b]Results for both NASHA/Dx and sham treatments ($n = 206$).

point) showed better correlations to secondary efficacy end points compared with the higher response thresholds (Table 2). The greatest correlations at the ≥ 40% and ≥ 50% response thresholds occurred with change from baseline in the number of FI episodes, the number of FI episodes when awake, and the number of FI-free days (r_s, 0.70–0.79). The strength of the correlation decreased with incremental increases in response threshold from 60% to 80%. The correlation coefficients for the change from baseline in the number of FI episodes with solid stools and with loose stools were greatest at the ≥ 40% and ≥ 50% response thresholds (r_s, 0.43–53) and tended to decrease at the higher response thresholds. The correlation between CCFIS total score and response threshold did not vary markedly across response thresholds (r_s, 0.36–0.42). Similarly, the correlation between scores for the coping and behavior, lifestyle, depression and self-perception, and embarrassment domains of the fecal incontinence quality of life (FIQOL) scale was low ($r_s < 0.3$) and did not vary markedly across response thresholds.

Based on the results of the correlational analyses, a responsiveness analysis was conducted using the ≥ 50% threshold employed in the study by Graf *et al.*[14] At this cutoff, responders (NASHA/Dx and sham combined) could reliably be differentiated from nonresponders (NASHA/Dx and sham combined) for all secondary end points ($P < 0.0001$ for all, except $P = 0.0003$ for the number of FI episodes when asleep and $P = 0.03$ for the number of controlled bowel-emptying episodes with fecal urgency; Table 3), with the exception of the number of controlled bowel-emptying episodes ($P = 0.48$).

Table 3 Responsiveness analysis at the ≥50% responder threshold[a]

Secondary end point, mean (s.d.)	Patients achieving ≥50% decrease ($n=92$)	Patients not achieving ≥50% decrease ($n=114$)	Difference	*P* value
CCFIS total score	−3.61 (3.46)	−1.08 (3.11)	−2.53	<0.0001
Total number of FI episodes	−18.3 (19.60)	−0.15 (10.62)	−18.19	<0.0001
Number of FI episodes when awake	−17.0 (18.49)	−0.31 (9.76)	−16.71	<0.0001
Number of FI episodes when asleep	−1.32 (3.01)	0.16 (2.80)	−1.48	0.0003
Number of FI episodes with solid stools	−8.96 (14.85)	1.22 (17.16)	−10.18	<0.0001
Number of FI episodes with loose stools	−9.38 (14.97)	−1.37 (18.85)	−8.01	<0.0001
Number of controlled bowel-emptying episodes with fecal urgency	−4.10 (11.23)	−0.79 (9.92)	−3.31	0.03
Number of controlled bowel-emptying episodes	0.86 (12.74)	0.47 (15.77)	0.39	0.49
Number of FI-free days	5.62 (3.08)	0.25 (2.68)	5.37	<0.0001

CCFIS, Cleveland Clinic Florida fecal incontinence score; FI, fecal incontinence; NASHA/Dx, nonanimal stabilized hyaluronic acid/dextranomer.
[a]NASHA/Dx and sham treatments combined.

DISCUSSION

Clinical studies of anal canal bulking agents have defined clinical response as a ≥50% decrease from baseline in the number of FI episodes.[11–14] The extent to which this threshold corresponds to clinically meaningful improvements has not been fully established.[11,12] Analyses of results of the study by Graf *et al.*[14] provide empirical data that not only support the use of the ≥50% threshold for treatment response in trials of FI but also describe the degree to which this threshold is valid and responsive.

In general in the current study, significant differences between the treatment groups were identified at thresholds from 40% through 80%. As expected, the percentage of patients achieving response in either treatment arm decreased as the threshold increased. An optimal threshold for treatment response should maintain the statistical power needed to accurately estimate differences between the treatment groups.[14,19] Increasing the threshold for response above what is clinically meaningful may decrease the statistical power needed to accurately identify treatment differences within the given sample size.[20] Further, setting a threshold for response below what is clinically meaningful has both ethical and financial implications, including unnecessarily exposing patients to study treatments that may provide little additional clinical benefit.[21] Given that the minimal clinically important difference for the primary outcome was calculated to be 35%, it was of interest to further examine the performance of response thresholds ranging from ≥40% to ≥80%. Correlational analyses and responsiveness analyses were conducted to identify the most valid and clinically meaningful responder threshold in patients treated with NASHA/Dx vs. sham injection.

Correlation analyses were conducted to help define the optimal response threshold with NASHA/Dx. Overall, the ≥40% and >50% response thresholds consistently had the greatest correlation with most secondary outcomes assessed. As the response threshold increased (i.e., from ≥60% to ≥80%), the strength of the association between treatment response and study outcome decreased. The primary efficacy outcome of this clinical study (i.e., a decrease from baseline of ≥50% in FI episodes)[14] correlated strongly with secondary efficacy end points of change from baseline in the number of FI episodes, number of FI episodes when the patient was awake, and the number of FI-free days.[22] Further, the correlation between the ≥40% and ≥50% response thresholds and a number of other secondary and exploratory efficacy end points were also still meaningful (i.e., $r \geq 0.4$).[23] However, there was no apparent correlation between FIQOL domain scores (i.e., coping and behavior, lifestyle, depression and self-perception, and embarrassment) and thresholds between 40% and 80%. The ability to evaluate correlations for FIQOL domain scores and various threshold levels may have been limited by the findings of Graf *et al.*,[14] which showed a significant difference relative to placebo in only the FIQOL coping and behavior domain score at 6 months. Comparing quality of life improvements to placebo rather than to within-subject baseline scores (the latter being an approach commonly used in studies of FI) may have hindered the ability to detect clear improvements in these domains.

The responsiveness analysis examined the ability of the ≥50% threshold to discriminate responders from nonresponders across clinically relevant symptoms known to be present in patients with FI.[19,20,24] Responders separated from nonresponders at the ≥50% threshold for all but one of the secondary end points evaluated (i.e., number of episodes of controlled bowel emptying episodes). Indeed, a number of clinical studies of NASHA/Dx for the treatment of FI have demonstrated a ≥50% decrease from baseline in the number of FI episodes.[11–14] Change from baseline in FI episodes and the number of FI-free days are variables that have been shown to significantly improve following treatment with NASHA/Dx and help to establish the ≥50% response threshold as clinically relevant.[12–14,20,24]

There are a few sham-controlled studies for treatments of FI, and the inclusion of a sham control group allowed for important comparisons to the NASHA/Dx group at different response thresholds. Although significant differences vs. sham injection were observed up to the 80% responder threshold (NASHA/Dx, 22.8%; sham, 10%; $P=0.02$), it appears that the 50% response threshold may be the one that is most valid and clinically relevant under the conditions tested.

This study has a number of limitations, including the *post hoc* nature of the data analyses. Another limitation is that the end points chosen for the analyses are the ones that were determined *a priori* in the original clinical study. FI is a multifactorial condition;[5] thus, inclusion of other outcomes (e.g., gas incontinence, use of fewer pads, complete resolution

of incontinence) may have yielded different results. Validation of these findings within other clinical studies of FI may provide further support for the use of the ≥50% response threshold.

In conclusion, this study validates the threshold of ≥50% reduction in FI episodes as a clinically relevant construct for patients with FI treated with NASHA/Dx. This finding is important for both the critical evaluation of the current data and the design of future clinical trials.

CONFLICT OF INTEREST

Guarantor of the article: Jaime E. Sanchez, MD, MSPH.
Specific author contributions: Interpretation of the data and drafting the manuscript: Jaime E. Sanchez; interpretation of the data, critically revising the manuscript, final approval: Darren M. Brenner and Howard Franklin; interpretation of the data, drafting, critically revising the manuscript, final approval: Jing Yu, Andrew C. Barrett, and Craig Paterson. All authors approved the final draft submitted.
Financial support: Funding for technical editorial and medical writing assistance was provided by Salix Pharmaceuticals Inc., Raleigh, NC, USA.
Potential competing interests: Dr Jaime E. Sanchez is a consultant and speaker for Salix Pharmaceuticals Inc., Dr Darren M. Brenner is a consultant and/or speaker for Salix Pharmaceuticals Inc., Ironwood Pharmaceuticals Inc., Forest Laboratories Inc., and Proctor & Gamble. Dr Howard Franklin, Dr Jing Yu, Dr Andrew C. Barrett, and Dr Craig Paterson are employees of and hold stock in Salix Pharmaceuticals Inc. This research was funded by Salix Pharmaceuticals Inc., Raleigh, NC, USA.

Acknowledgments. Technical editorial and medical writing assistance was provided, under the direction of the authors, by Sophie Bolick, PhD, Synchrony Medical Communications, LLC, West Chester, PA, USA. Funding for this research was provided by Salix Pharmaceuticals Inc., Raleigh, NC, USA.

Study Highlights

WHAT IS CURRENT KNOWLEDGE

- ✓ Definitions of clinical response in studies of fecal incontinence treatments vary widely.
- ✓ A single accepted response threshold for treatment with NASHA/Dx has not been established.

WHAT IS NEW HERE

- ✓ A responder threshold of ≥50% reduction from baseline in fecal incontinence episodes appears valid and responsive.

1. Tan JJ, Chan M, Tjandra JJ. Evolving therapy for fecal incontinence. *Dis Colon Rectum* 2007; **50**: 1950–1967.
2. Ditah I, Devaki P, Luma HN *et al.* Prevalence, trends, and risk factors for fecal incontinence in United States adults, 2005-2010. *Clin Gastroenterol Hepatol* 2014; **12**: 636–643.
3. Whitehead WE, Borrud L, Goode PS *et al.* Fecal incontinence in US adults: epidemiology and risk factors. *Gastroenterology* 2009; **137**: 512–517, 517.
4. Rao SS. Diagnosis and management of fecal incontinence. American College of Gastroenterology Practice Parameters Committee. *Am J Gastroenterol* 2004; **99**: 1585–1604.
5. Bharucha AE. Fecal incontinence. *Gastroenterology* 2003; **124**: 1672–1685.
6. Rockwood TH, Church JM, Fleshman JW *et al.* Fecal Incontinence Quality of Life Scale: quality of life instrument for patients with fecal incontinence. *Dis Colon Rectum* 2000; **43**: 9–16.
7. Shah BJ, Chokhavatia S, Rose S. Fecal incontinence in the elderly: FAQ. *Am J Gastroenterol* 2012; **107**: 1635–1646.
8. Van Koughnett JA, Wexner SD. Current management of fecal incontinence: choosing amongst treatment options to optimize outcomes. *World J Gastroenterol* 2013; **19**: 9216–9230.
9. Vaizey C. Faecal incontinence: standardising clinical trials. *Colorectal Dis* 2014; **16**: 156–158.
10. Jensen LL, Lowry AC. Biofeedback improves functional outcome after sphincteroplasty. *Dis Colon Rectum* 1997; **40**: 197–200.
11. La Torre F, de la Portilla F. Long-term efficacy of dextranomer in stabilized hyaluronic acid (NASHA/Dx) for treatment of faecal incontinence. *Colorectal Dis* 2013; **15**: 569–574.
12. Dodi G, Jongen J, de la Portilla F *et al.* An open-label, noncomparative, multicenter study to evaluate efficacy and safety of NASHA/Dx gel as a bulking agent for the treatment of fecal incontinence. *Gastroenterol Res Pract* 2010; **2010**: 467136.
13. Danielson J, Karlbom U, Sonesson AC *et al.* Submucosal injection of stabilized nonanimal hyaluronic acid with dextranomer: a new treatment option for fecal incontinence. *Dis Colon Rectum* 2009; **52**: 1101–1106.
14. Graf W, Mellgren A, Matzel KE *et al.* Efficacy of dextranomer in stabilised hyaluronic acid for treatment of faecal incontinence: a randomised, sham-controlled trial. *Lancet* 2011; **377**: 997–1003.
15. de Groot IB, Favejee MM, Reijman M *et al.* The Dutch version of the knee injury and osteoarthritis outcome score: a validation study. *Health Qual Life Outcomes* 2008; **6**: 16.
16. Alexander M, Berger W, Buchholz P *et al.* The reliability, validity, and preliminary responsiveness of the Eye Allergy Patient Impact Questionnaire (EAPIQ). *Health Qual Life Outcomes* 2005; **3**: 67.
17. Giesinger JM, Kuster MS, Behrend H *et al.* Association of psychological status and patient-reported physical outcome measures in joint arthroplasty: a lack of divergent validity. *Health Qual Life Outcomes* 2013; **11**: 64.
18. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: the remarkable universality of half a standard deviation. *Med Care* 2003; **41**: 582–592.
19. Irvine EJ, Whitehead WE, Chey WD *et al.* Design of treatment trials for functional gastrointestinal disorders. *Gastroenterology* 2006; **130**: 1538–1551.
20. Snapinn SM, Jiang Q. Responder analyses and the assessment of a clinically relevant treatment effect. *Trials* 2007; **8**: 31.
21. Altman DG. Statistics and ethics in medical research: III How large a sample? *Br Med J* 1980; **281**: 1336–1338.
22. Taylor R. Interpretation of the correlation coefficient: a basic review. *J Diagn Med Sonogr* 1990; **6**: 35–39.
23. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; **33**: 159–174.
24. Uryniak T, Chan ISF, Fedorov VV *et al.* Responder analyses-a PhRMA position paper. *Stat Biopharm Res* 2011; **3**: 476–487.