AP&T Alimentary Pharmacology and Therapeutics

Psychometric validation of symptom severity measures in irritable bowel syndrome with constipation

V. S. L. Williams*, L. M. Nelson*, S. E. Fehnel*, J. MacDougall†, R. T. Carson‡, S. Tourkodimitris‡, C. Kurtz†, M. J. Baird† & J. M. Johnston†

*RTI Health Solutions, Durham, NC, USA.
†Ironwood Pharmaceuticals, Cambridge, MA, USA.
‡Forest Research Institute, Jersey City, NJ, USA.

Correspondence to:
Dr V. Williams, 3040 Cornwallis Road, PO Box 12194, Research Triangle Park, NC 27709, USA.
E-mail: vwilliams@rti.org

Publication data
Submitted 16 December 2013
First decision 20 January 2014
Resubmitted 31 March 2014
Resubmitted 20 May 2014
Accepted 20 May 2014
EV Pub Online 18 June 2014

This article was accepted for publication after full peer-review.

SUMMARY

Background
Historically, measures of symptom severity of irritable bowel syndrome with constipation (IBS-C) in clinical trials have not met the evidence requirements described in the FDA guidance on patient-reported outcomes (PROs), which describes the evidentiary requirements and review criteria for patient-reported outcome measures intended to support product approval or labelling claims.

Aim
Data from two phase 3 trials (N = 1608) of linaclotide for the treatment of IBS-C were analysed to evaluate the psychometric properties of patient-reported outcome measures assessing changes in the severity of abdominal and bowel symptoms.

Methods
A set of patient-reported outcome assessments addressing abdominal and bowel symptoms, the IBS-C Symptom Severity Measures, were administered daily using interactive voice response system technology. Intraclass correlation coefficients (ICCs), Pearson correlations, factor analyses, F-tests and effect sizes were computed to evaluate the reliability, construct validity, discriminating ability and responsiveness of the IBS-C Symptom Severity Measures in a clinical trial context.

Results
The IBS-C Symptom Severity Measures showed highly satisfactory test–retest reliability (ICCs ranging from 0.79 to 0.95) and construct validity. Factor analyses indicated one factor for abdominal symptoms and another for bowel symptoms. Known-groups F-tests comparing subgroups based on various responder definitions were statistically significant and in the expected direction, substantiating the discriminating ability of the IBS-C Symptom Severity Measures. Responsiveness statistics (ranging from 0.6 to 2.1) demonstrated these measures are also capable of detecting change.

Conclusions
The psychometric analysis results strongly support the reliability, construct validity, discriminating ability and responsiveness of the IBS-C Symptom Severity Measures and substantiate the conclusion of linaclotide treatment benefit.

Aliment Pharmacol Ther 2014; **40**: 298–308

doi:10.1111/apt.12830

INTRODUCTION

Irritable bowel syndrome with constipation (IBS-C) is a functional bowel disorder primarily characterised by patient reports of recurrent abdominal pain and/or discomfort, accompanied by altered bowel function.[1, 2] IBS prevalence estimates range from 3% to 20% of North Americans, with IBS-C accounting for approximately one third of all IBS cases.[1, 3, 4] According to Lovell and colleagues,[5] the estimated global prevalence of IBS is 11.2%.

IBS-C is a symptomatic disorder; therefore, the primary aim of therapeutic intervention is to improve IBS symptoms important and bothersome to patients. Reliable and valid patient-reported outcome (PRO) measures are needed to directly assess patient's response to therapy. Based on a synthesis of relevant literature and patient interviews, and in accordance with recommendations set forth in the Food and Drug Administration's (FDA's) guidance, *Patient-Reported Outcome (PRO) Measures: Use in Medical Product Development to Support Labeling Claims*,[6] a novel set of patient-reported IBS-C Symptom Severity Measures was developed to assess abdominal and bowel symptoms within the linaclotide clinical development programme.[7] Based on these PRO symptom measures, primary and secondary endpoints were developed for a phase 2b study and two phase 3 trials investigating the efficacy and safety of linaclotide, a minimally absorbed peptide agonist of the intestinal guanylate cyclase type-C receptor, in patients with IBS-C.[8–10]

The objective of the present study was to evaluate and document the psychometric or measurement properties of the newly developed patient-reported IBS-C Symptom Severity Measures as recommended in the FDA's PRO guidance.

METHODS

Study design

Data from two randomised, multicentre, double-blind, placebo-controlled, parallel-group phase 3 clinical trials comparing placebo and oral linaclotide (290 μg) taken once daily were used to evaluate the psychometric properties of the bowel and abdominal symptoms of the IBS-C Symptom Severity Measures. The two clinical trials included 803 and 805 adult patients with IBS-C; both trials included a 2-week pre-treatment period followed by 12 weeks of treatment. One of the trials included an additional 4-week randomised withdrawal period following the 12-week treatment period, while the other trial included an additional 14 weeks of treatment, totalling 26 weeks of treatment. In brief, key inclusion criteria were as follows: patients 18 years or older who met the Rome II criteria for IBS,[11] and who reported fewer than three spontaneous bowel movements (SBMs) per week, along with one or more of the following symptoms for at least 12 weeks in the preceding 12 months: (i) straining during >25% of bowel movements (BMs); (ii) lumpy or hard stools during >25% of BMs or (iii) sensation of incomplete evaluation during >25% of BMs. In addition, during the 2-week pre-treatment period, patients were required to have a mean score ≥3 for abdominal pain at its worst on an 11-point numerical rating scale (NRS) ranging from 0 = no abdominal pain to 10 = very severe abdominal pain, as well as an average of <3 complete SBMs (CSBMs) and ≤5 SBMs per week. Further details are provided in Chey *et al.* [9] and Rao *et al.*[10]

This study was reviewed and approved by the appropriate ethics committees at the participating centres and approvals were obtained prior to any subject's participation. All study subjects provided written informed consent.

Measures

IBS-C symptom severity measures. Participant responses to the IBS-C Symptom Severity Measures were gathered via telephone-based interactive voice response system (IVRS) technology. During a daily IVRS telephone call, patients were asked a series of questions pertaining to the frequency and timing of their BMs and their use of rescue medication. For each BM, patients were asked about completeness of evacuation and straining, and to describe stool consistency using the Bristol Stool Form Scale.[12] Five daily measures, Abdominal Pain, Abdominal Discomfort, Bloating, Abdominal Fullness and Abdominal Cramping were administered by IVRS using an 11-point NRS.

Using the daily IVRS bowel information, weekly frequencies of SBMs (SBM Frequency) and CSBMs (CSBM Frequency) were calculated. Average weekly severity was computed for the two BM-specific items, Straining and Stool Consistency, as well as for the five daily Abdominal Pain (at its worst), Abdominal Discomfort, Bloating, Abdominal Fullness and Abdominal Cramping items. Baseline values were computed as the average of the two pre-treatment weeks.

Patient ratings of constipation and IBS symptom severity. Two items administered weekly via IVRS were used in the evaluation of the IBS-C Symptom Severity

Measures: Constipation Severity and IBS Symptom Severity during the past 7 days. For both of these global items, patients rated their symptoms on a 5-point ordered response scale ranging from none to very severe.

Patient ratings of change. Patient rating of change questions (PRCQs) for each of the symptoms addressed by the IBS-C Symptom Severity Measures were asked at all trial visits during the treatment period. For example, to assess change in abdominal pain, patients were asked to rate their relief of abdominal pain at its worst during the past 7 days compared to before the start of the trial (PRCQ Abdominal Pain). The PRCQ Degree of IBS Relief, which asked patients to compare their IBS symptoms during the past 7 days to their symptoms prior to the start of the study, was collected weekly via IVRS. For all symptom-specific PRCQs and PRCQ Degree of IBS Relief, patients responded using a 7-point balanced response scale (1 = completely relieved/improved, 2 = considerably relieved/improved, 3 = somewhat relieved/improved, 4 = unchanged, 5 = somewhat worse, 6 = considerably worse and 7 = as bad as I can imagine). Another item administered weekly, Adequate Relief, asked patients if they experienced adequate relief from their IBS symptoms during the past 7 days (1 = Yes, 2 = No).

Trial Endpoints. Subgroups for evaluating the validity and responsiveness of the IBS-C Symptom Severity Measures were defined using selected trial endpoints: (1) the 9/12 CSBM 3 + 1 Responder, defined as patients who report ≥ 3 CSBMs per week and an increase of ≥ 1 CSBM per week (in the same week) compared with baseline (weekly CSBM 3 + 1 Responder) for at least 9 of 12 weeks of the treatment period; (2) the FDA Responder for IBS-C,[13] defined as patients who report an improvement of ≥ 30% more than baseline in the average of the daily worst Abdominal Pain score (weekly Abdominal Pain Responder) and report an increase of ≥ 1 CSBM per week more than baseline (weekly CSBM +1 Responder), in the same week for at least 6 of 12 weeks; and (3) the 9/12 IBS Symptom Severity Responder, defined as patients with a decrease of ≥ 1 in the mean IBS Symptom Severity score from baseline for at least 9 of 12 weeks. Notably, the last endpoint is particularly suitable for the psychometric evaluation because it is independent of specific abdominal and bowel symptoms, and requires patients to reflect on their current condition.

IBS-QOL, IBS-SSS and HADS. The following measures were administered at the beginning and end of treatment: (1) the Irritable Bowel Syndrome Quality of Life (IBS-QOL), a 34-item questionnaire that assesses the degree to which IBS interferes with quality of life (QOL)[14] and (2) the Irritable Bowel Syndrome Symptom Severity Scale (IBS-SSS), a 5-item questionnaire that measures the severity and frequency of abdominal pain, the severity of abdominal distention, dissatisfaction with bowel habits and interference with QOL.[15] IBS-QOL scores range from 0 to 100 with higher scores indicating better QOL; the IBS-SSS ranges from 0 to 500 with higher scores indicating worse disease severity and impact. The Hospital Anxiety and Depression Scale (HADS)[16, 17] was administered once on the first day of treatment. The HADS is a 14-item self-report measure designed to screen for anxiety (HADS-A) and depression (HADS-D) in medically ill patients. Both HADS scores range from 0 to 21 with higher values reflecting greater severity.

Analytic methods

All psychometric properties were investigated irrespective of treatment group using SAS for Windows version 9.2.[18] Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were conducted using Mplus version 6.1.[19]

Reliability. To assess test–retest reliability, intraclass correlation coefficients (ICCs) were computed for each of the IBS-C Symptom Severity Measures using data from the last 2 weeks of treatment when patients' symptoms were assumed to be stable based on findings from the phase 2b trials of linaclotide.[20] The ICCs were computed using all available (i.e. nonmissing) data from weeks 11 ('test') and 12 ('retest'). It is generally recommended that ICCs be at least 0.70 for multi-item scales (e.g. Nunnally and Bernstein [21]).

Validity. A number of specific interitem relationships between the IBS-C Symptom Severity Measures were hypothesised *a priori*. Overall, abdominal symptom measures were anticipated to intercorrelate more strongly than were bowel symptom intercorrelations, based on the results of qualitative research conducted with patients. Cohen (1988) provides a general rule-of-thumb for the interpretation of correlation coefficients: correlations of about 0.10 are small, those of about 0.30 are medium or moderate, and those of about 0.50 or greater are considered large.[22] In addition, evidence of construct validity was based on the correlations between change from baseline in the IBS-C Symptom Severity Measures

and the corresponding PRCQ ($|r| > 0.50$), as well as the weekly Degree of IBS Relief item. Abdominal Pain, Abdominal Discomfort and Bloating items were expected to correlate highly ($|r| > 0.50$) with the IBS-SSS total score. However, because health-related QOL is a more distal concept from symptoms, the IBS-C Symptom Severity Measures were anticipated to exhibit relatively lower correlations (i.e. $|r| < 0.30$) with the IBS-QOL Overall score and the eight IBS-QOL subscale scores. Low correlations were anticipated between the IBS-C Symptom Severity Measures and the HADS-A and HADS-D because these are not IBS-specific measures.

Data from the phase 3 trials were pooled and randomly divided into two approximately equal subsamples (with the randomisation stratified by trial). EFA was performed on one subsample to determine the number of continuous latent variables (i.e. factors) that optimally explain the correlations observed among the IBS-C symptom severity measures. We anticipated a two-dimensional structure, with abdominal symptom measures loading on one factor and the bowel symptom measures on the other factor. Models were proposed based on the EFA results and evaluated using CFA methods and data from the remaining sample.

The discriminating ability of the IBS-C Symptom Severity Measures was evaluated using analyses of variance (ANOVAS) to examine mean differences in the IBS-C Symptom Severity Measure scores between patients classified into groups based on the 9/12 CSBM 3 + 1 Responder endpoint, the FDA Responder endpoint and the IBS Symptom Severity Responder endpoint. It was predicted that scores would be statistically better among patients classified as responders, thereby providing evidence in support of the discriminating ability of the IBS-C Symptom Severity Measures.

Responsiveness. To evaluate responsiveness, a variant of Guyatt's statistic[23] was calculated to compare responders to nonresponders based on the 9/12 CSBM 3 + 1 Responder, the FDA Responder and the IBS Symptom Severity Responder endpoint definitions. Computing change as the difference between baseline and the average change across the treatment period, a responsiveness effect size was calculated for the IBS-C Symptom Severity Measures as the difference in mean item change in responders and nonresponders divided by the standard deviation (s.d.) of the mean item change in the nonresponders. The resulting value is an indication of a measure's ability to detect change in IBS-C symptoms. Effect sizes of about 0.20 represent small effects, those of about 0.50 represent moderate effects, and those ≥ 0.80 represent large effects.[22]

RESULTS

The majority of participants in the phase 3 trials were female (over 90%) and White (over 75%), with an average age of approximately 44 years, ranging from 18 to 87 years.[8–10] Table 1 summarises the baseline scores on the IBS-C Symptom Severity Measures.

Reliability

The test–retest ICCs for bowel symptoms exceeded the recommended 0.70, ranging from 0.79 for Stool Consistency (n = 1194) to 0.86 for SBM Frequency and CSBM Frequency (n = 1602). ICCs for the abdominal symptoms were even higher, ranging from 0.94 (Abdominal Cramping, n = 1284) to 0.95 (Abdominal Pain, Abdominal Discomfort and Bloating, n = 1602; Abdominal Fullness, n = 1284).

Table 1 | Baseline descriptive statistics

IBS-C symptom severity measure (N = 1602)	Mean (s.d.)
CSBM Frequency	0.2 (0.5)
SBM Frequency	1.8 (1.4)
Stool Consistency	2.3 (1.0)
Straining	3.5 (0.8)
Abdominal Pain	5.6 (1.7)
Abdominal Discomfort	6.1 (1.7)
Bloating	6.6 (1.8)
Abdominal Cramping	5.3 (1.9)
Abdominal Fullness	6.6 (1.8)
Additional PRO measures (N = 1595–1601)	Mean (s.d.)
IBS-SSS total	364.2 (74.5)
HADS anxiety	7.6 (3.9)
HADS depression	3.9 (3.4)
IBS-QOL: overall	61.1 (20.7)
IBS-QOL: dysphoria	62.3 (24.7)
IBS-QOL: interference with activity	68.0 (23.2)
IBS-QOL: body image	48.2 (22.8)
IBS-QOL: health worry	44.9 (25.0)
IBS-QOL: food avoidance	50.1 (29.8)
IBS-QOL: social reaction	67.1 (25.3)
IBS-QOL: sexual	67.9 (31.5)
IBS-QOL: relationships	73.4 (23.8)

CSBM, complete spontaneous bowel movement; HADS, Hospital Anxiety and Depression Scale; IBS-C, irritable bowel syndrome with constipation; IBS-SSS, irritable bowel symptom severity scale; IBS-QOL, irritable bowel syndrome quality of life questionnaire; SBM, spontaneous bowel movement; s.d., standard deviation.

Validity

Construct validity. Table 2 presents the interitem correlations among the IBS-C Symptom Severity Measures computed using treatment period averages. As expected based on qualitative data collected from patients, Bloating correlated more highly with Abdominal Discomfort than with Abdominal Pain. In addition, as predicted, Abdominal Discomfort and Abdominal Pain correlated strongly. Notably, Abdominal Pain, a primary endpoint, was strongly correlated with the other abdominal symptom measures and moderately correlated with the bowel symptom measures. CSBM Frequency, also a primary endpoint, correlated strongly with SBM Frequency and Straining, and moderately with Stool Consistency as well as the abdominal symptom measures.

As expected, Constipation Severity correlated strongly with CSBM and SBM Frequency and Straining, while the correlation of Constipation Severity with Stool Consistency was moderate. Strong correlations were hypothesised between Constipation Severity and the abdominal symptom measures, and in fact, strong correlations were observed.

Similar to Constipation Severity, IBS Symptom Severity was expected to correlate strongly with CSBM Frequency, SBM Frequency, Stool Consistency and Straining. As shown in Table 2, IBS Symptom Severity correlated strongly with Straining and CSBM Frequency, and moderately with SBM Frequency and Stool Consistency. As hypothesised, strong correlations were observed between IBS Symptom Severity and the abdominal symptom measures. Correlations were moderate to strong as expected between Adequate Relief and the IBS-C Symptom Severity Measures.

Table 3 displays the correlations between the weekly assessment of Degree of IBS Relief, which asked patients to compare their baseline symptoms to those at each treatment week, and corresponding changes from baseline in the IBS-C Symptom Severity Measures. The correlations associated with CSBM Frequency, SBM Frequency and Stool Consistency were negative because lower values for these variables denote worse outcomes; for the rest of the IBS-C Symptom Severity Measures, higher values indicate worse outcomes. All correlations were at least moderate in size and statistically significant.

Table 2 | Interitem correlations among IBS-C symptom severity measures (treatment period averages) (N = 1531–1602)

IBS-C symptom severity measure	CSBM Frequency	SBM Frequency	Stool Consistency	Straining	Abdominal Pain	Abdominal Discomfort	Bloating	Abdominal Cramping	Abdominal Fullness
CSBM Frequency	–								
SBM Frequency	0.69	–							
Stool Consistency	0.43	0.48	–						
Straining	−0.52	−0.44	−0.62	–					
Abdominal Pain	−0.40	−0.30	−0.26	0.49	–				
Abdominal Discomfort	−0.43	−0.31	−0.26	0.51	0.93	–			
Bloating	−0.44	−0.31	−0.25	0.50	0.81	0.90	–		
Abdominal Cramping	−0.36	−0.28	−0.25	0.47	0.91	0.85	0.74	–	
Abdominal Fullness	−0.44	−0.31	−0.25	0.50	0.80	0.89	0.96	0.73	–
Constipation severity	−0.61	−0.56	−0.49	0.71	0.59	0.64	0.63	0.52	0.64
IBS symptom severity	−0.53	−0.44	−0.41	0.67	0.65	0.70	0.68	0.58	0.67
Adequate relief	0.52	0.45	0.42	−0.54	−0.49	−0.55	−0.51	−0.42	−0.52

All correlations are statistically significant, $P < 0.01$.

CSBM, complete spontaneous bowel movement; IBS-C, irritable bowel syndrome with constipation; SBM, spontaneous bowel movement.

Aliment Pharmacol Ther 2014; 40: 298-308

Table 3 | Correlations between degree of IBS relief and changes in IBS-C symptom severity measures

IBS-C symptom severity measure	BL to week 1	BL to week 2	BL to week 4	BL to week 8	BL to week 12	BL to treatment average
CSBM Frequency	−0.41	−0.45	−0.49	−0.45	−0.48	−0.54
SBM Frequency	−0.40	−0.43	−0.44	−0.41	−0.42	−0.48
Stool Consistency	−0.28	−0.33	−0.34	−0.33	−0.37	−0.39
Straining	0.33	0.41	0.46	0.44	0.49	0.49
Abdominal Pain	0.40	0.47	0.54	0.54	0.56	0.61
Abdominal Discomfort	0.42	0.51	0.57	0.56	0.60	0.63
Bloating	0.45	0.53	0.56	0.56	0.58	0.63
Abdominal Cramping	0.39	0.47	0.52	0.52	0.55	0.59
Abdominal Fullness	0.46	0.55	0.57	0.57	0.60	0.64

BL, baseline; CSBM, complete spontaneous bowel movement; SBM, spontaneous bowel movement.

All correlations are statistically significant, $P < 0.01$.

Degree of IBS Relief: 1 = completely relieved/improved, 2 = considerably relieved/improved, 3 = somewhat relieved/improved, 4 = unchanged, 5 = somewhat worse, 6 = considerably worse and 7 = as bad as I can imagine.

The correlations between Degree of IBS Relief and change in Abdominal Pain, Abdominal Discomfort, Bloating, Abdominal Cramping and Abdominal Fullness are the highest using change from baseline to treatment period average, as opposed to change from baseline to individual study visit weeks.

The correlations between changes in each IBS-C Symptom Severity Measure and the corresponding symptom-related PRCQ further support the convergent validity of the IBS-C Symptom Severity Measures (data not shown). The majority of these correlations were at least moderate in size, statistically significant, and consistent with the correlations between the IBS-C Symptom Severity Measures and Degree of IBS Relief. Specifically, no correlation between any given IBS-C Symptom Severity Measure and the corresponding PRCQ differed by more than ±0.08 from the correlation between that IBS-C Symptom Severity Measure and Degree of IBS Relief (Table 3).

The IBS-SSS total score was strongly correlated with the abdominal symptom measures at baseline: Abdominal Pain $r = 0.60$, Abdominal Discomfort $r = 0.60$, Bloating $r = 0.57$, Abdominal Cramping $r = 0.54$ and Abdominal Fullness $r = 0.56$. Likely because of symptom improvement among responders resulting in increased variability in scores, these correlations increased during the last 4 weeks of the treatment period (Abdominal Pain $r = 0.68$, Abdominal Discomfort $r = 0.71$, Bloating $r = 0.68$, Abdominal Cramping $r = 0.62$ and Abdominal Fullness $r = 0.70$). Table 4 displays the correlations between the IBS-C Symptom Severity Measures and the IBS-QOL. As hypothesised, the correlations between the IBS-C Symptom Severity Measures and the IBS-QOL Overall and subscale scores were generally lower (i.e. $|r| < 0.30$) at baseline than towards the end of the treatment period. However, correlations computed using data averaged over the last 4 weeks of the treatment period were greater than expected, with moderate associations observed between the IBS-QOL Overall and Straining, Abdominal Pain, Abdominal Discomfort, Bloating, Abdominal Cramping and Abdominal Fullness. Also, as hypothesised, the IBS-C Symptom Severity Measures were weakly correlated with HADS-A (range −0.03 with CSBM Frequency to 0.09 with Abdominal Fullness) and HADS-D (range −0.06 with CSBM Frequency to 0.13 with Abdominal Discomfort) scores.

Factor analysis. The EFA of one random subsample of the pooled treatment period data indicated that two factors could be extracted, and the two-factor EFA produced the most interpretable and parsimonious solution, consistent with the factor analytic findings based on the phase 2b trials of linaclotide: One factor consisted of abdominal symptoms (Abdominal Pain, Abdominal Discomfort, Bloating, Abdominal Cramping and Abdominal Fullness), while the other factor consisted of bowel symptoms (SBM Frequency, CSBM Frequency, Stool Consistency and Straining).

The solution from the two-factor EFA model was evaluated with CFA using the remaining random subsample of the pooled treatment period data. Five pairs of measures – CSBM Frequency and SBM Frequency, Stool Consistency and Straining, Abdominal Pain and

Table 4 | Validity correlations between IBS-C symptom severity measures and IBS-QOL subscale scores at baseline and last 4 weeks

IBS-C Symptom Severity Measure	IBS-QOL overall	Dysphoria	Interference with activity	Body image	Health worry	Food avoidance	Social reaction	Sexual	Relationships
CSBM Frequency – BL	0.13*	0.11*	0.08*	0.17*	0.17*	0.08*	0.10*	0.12*	0.07*
Last 4 weeks	0.25*	0.23*	0.14*	0.31*	0.27*	0.22*	0.19*	0.19*	0.17*
SBM Frequency – BL	0.07*	0.06	0.03	0.08*	0.11*	0.01	0.06	0.05	0.04
Last 4 weeks	0.15*	0.14*	0.02	0.22*	0.22*	0.15*	0.13*	0.12*	0.11*
Stool Consistency – BL	0.03	0.02	0.03	0.02	0.04	0.01	0.02	0.05	0.03
Last 4 weeks	0.14*	0.11*	0.03	0.19*	0.16*	0.16*	0.10*	0.11*	0.12*
Straining – BL	−0.21*	−0.17*	−0.18*	−0.16*	−0.19*	−0.13*	−0.21*	−0.13*	−0.17*
Last 4 weeks	−0.35*	−0.31*	−0.24*	−0.38*	−0.35*	−0.30*	−0.28*	−0.23*	−0.25*
Abdominal Pain – BL	−0.31*	−0.29*	−0.29*	−0.25*	−0.21*	−0.24*	−0.27*	−0.19*	−0.24*
Last 4 weeks	−0.35*	−0.32*	−0.29*	−0.36*	−0.34*	−0.28*	−0.28*	−0.25*	−0.27*
Abdominal Discomfort – BL	−0.34*	−0.32*	−0.28*	−0.32*	−0.24*	−0.26*	−0.27*	−0.22*	−0.24*
Last 4 weeks	−0.38*	−0.35*	−0.30*	−0.42*	−0.37*	−0.31*	−0.30*	−0.27*	−0.28*
Bloating – BL	−0.29*	−0.27*	−0.20*	−0.39*	−0.24*	−0.24*	−0.22*	−0.19*	−0.17*
Last 4 weeks	−0.38*	−0.34*	−0.27*	−0.48*	−0.37*	−0.30*	−0.29*	−0.25*	−0.25*
Abdominal Cramping – BL	−0.30*	−0.26*	−0.30*	−0.24*	−0.21*	−0.21*	−0.26*	−0.18*	−0.24*
Last 4 weeks	−0.31*	−0.27*	−0.27*	−0.32*	−0.30*	−0.25*	−0.25*	−0.22*	−0.25*
Abdominal Fullness – BL	−0.30*	−0.29*	−0.22*	−0.37*	−0.26*	−0.24*	−0.23*	−0.19*	−0.17*
Last 4 weeks	−0.37*	−0.34*	−0.27*	−0.45*	−0.38*	−0.31*	−0.29*	−0.25*	−0.25*

BL, baseline; CSBM, complete spontaneous bowel movement; IBS-C, irritable bowel syndrome with constipation; IBS-QOL, irritable bowel syndrome quality of life questionnaire; SBM, spontaneous bowel movement.

The IBS-QOL reference period was over the past month. Because the IBS-QOL was administered at the last visit of the 12-week treatment period, the IBS-C Symptom Severity Items were averaged over weeks 9, 10, 11 and 12.

* $P < 0.01$.

Bloating, Abdominal Pain and Abdominal Cramping, and Bloating and Abdominal Fullness – were highly correlated, and when their errors were allowed to covary, the model fit was considerably improved. The CFA factor loadings are displayed in Figure 1.

Discriminating ability. It was predicted that IBS-C Symptom Severity Measures scores would be statistically ($P < 0.01$) better among patients classified as responders, and in fact, all tests of mean differences were statistically significant and in the predicted direction. Figure 2 displays

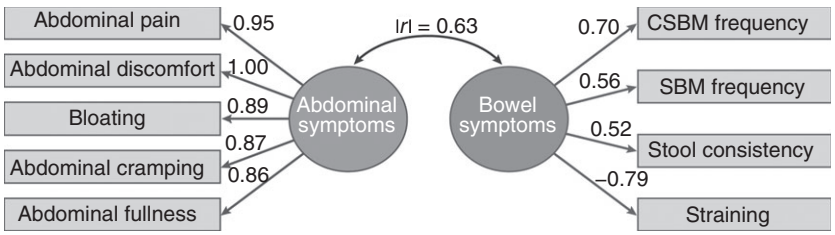


Figure 1 | Illustration of the 2-factor confirmatory factor analysis model for IBS-C symptom severity measures, with factor loadings. IBS-C, irritable bowel syndrome with constipation; CSBM, complete spontaneous bowel movement; SBM, spontaneous bowel movement. The CFA solution in Figure 1 includes correlated residuals that are not shown between the following: CSBM Frequency and SBM Frequency; Stool Consistency and Straining; Abdominal Pain and Bloating; Abdominal Pain and Abdominal Cramping; Bloating and Abdominal Fullness.

Aliment Pharmacol Ther 2014; 40: 298-308

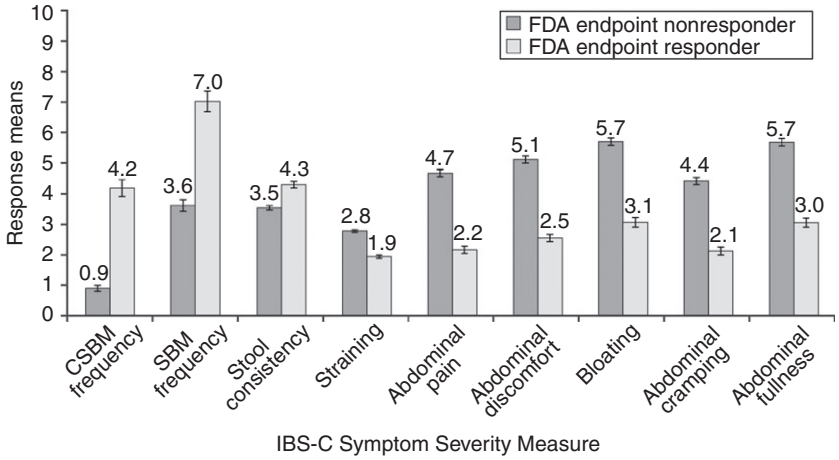
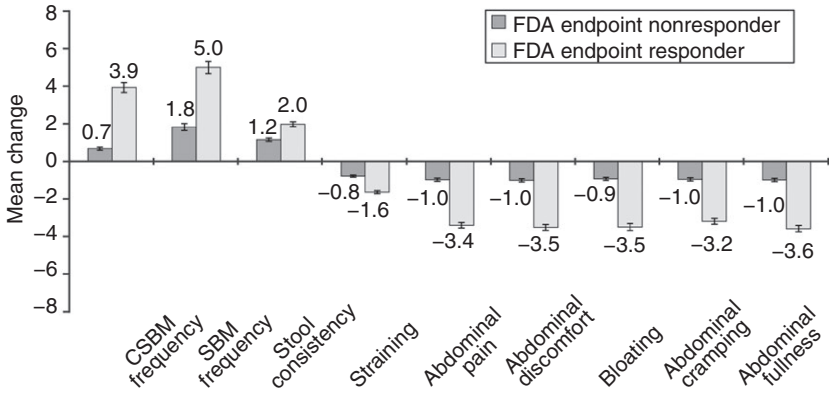


Figure 2 | Known-groups comparison – mean scores on IBS-C symptom severity measures at week 12 for subgroups based on the FDA endpoint responder vs. nonresponder. CSBM, complete spontaneous bowel movement; FDA, Food and Drug Administration; IBS-C, irritable bowel syndrome with constipation; SBM, spontaneous bowel movement. Higher mean scores for CSBM Frequency, SBM Frequency and Stool Consistency indicate better outcomes, whereas lower mean scores for Straining, Abdominal Pain, Abdominal Discomfort, Bloating, Abdominal Cramping and Abdominal Fullness indicate better outcomes.

the responder and nonresponder subgroup means for each of the IBS-C Symptom Severity Measures based on the FDA's recommended responder threshold. These results demonstrate that the IBS-C Symptom Severity Measures are capable of distinguishing between study participants classified according to a variety of responder criteria.

Responsiveness

Figure 3 displays the mean changes and responsiveness effect sizes based on the FDA's recommended responder threshold. This responsiveness effect size, a standardised change score in units based on the s.d. of change in the nonresponder group, was large for all of the IBS-C



Effect size	2.1	1.0	0.6	−1.0	−1.6	−1.7	−1.8	−1.5	−1.8

Figure 3 | Responsiveness – effect sizes and mean change in scores from baseline to week 12 on IBS-C symptom severity measures for subgroups based on the FDA endpoint responder vs. nonresponder. CSBM, complete spontaneous bowel movement; FDA, Food and Drug Administration; IBS-C, irritable bowel syndrome with constipation; SBM, spontaneous bowel movement. Positive effect sizes and mean changes in scores for CSBM Frequency, SBM Frequency and Stool Consistency indicate better outcomes, whereas negative effect sizes and mean changes in scores for straining, abdominal pain, abdominal discomfort, bloating, abdominal cramping and abdominal fullness indicate better outcomes.

Symptom Severity Measures, with the one exception of moderate effects for Stool Consistency.

DISCUSSION

A comprehensive set of IBS-C PRO symptom measures was developed for the IBS-C linaclotide trials in accordance with the evidentiary standards outlined in the FDA's PRO guidance.[6] These measures were used to develop primary and secondary endpoints in two phase 3 trials demonstrating the efficacy and safety of linaclotide in patients with IBS-C.[8–10] The present study evaluated the psychometric properties of reliability, validity and responsiveness of the set of measures assessing bowel symptoms and abdominal symptoms using pooled data from the two linaclotide phase 3 IBS-C trials.

Overall, the psychometric properties of the IBS-C Symptom Severity Measures were very satisfactory. Test–retest reliability was acceptable and evidence for construct validity was supported for both abdominal symptom measures (Abdominal Pain, Abdominal Discomfort, Bloating, Abdominal Cramping and Abdominal Fullness) and bowel symptom measures (SBM Frequency, CSBM Frequency, Stool Consistency and Straining). Predicted interitem correlations among the IBS-C Symptom Severity Measures were as expected and were consistent with the patterns and trends observed in the phase 2b study. Notably, CSBM Frequency and Abdominal Pain, which are recommended by the FDA to define primary endpoints in IBS-C trials and are required to support drug approval,[13] were correlated moderately to strongly with the other bowel and abdominal symptom measures. There were also some extremely large correlations, for example, between Abdominal Fullness and Bloating, Abdominal Discomfort and Abdominal Pain, and Abdominal Cramping and Abdominal Pain. Such strong correlations indicate redundancy and in the context of multi-item scale construction would normally afford the opportunity for item deletion. However, decisions about item reduction must be based on the totality of evidence gathered – both quantitative and qualitative results. The PRO Guidance recommends that all symptoms important to patients be measured to assure content validity and IBS-C patients reported that these were all important components of IBS-C.

The strength of the correlations between the change in each IBS-C Symptom Severity Measure and Degree of IBS Relief and the symptom-specific PRCQs at corresponding time points was moderate to strong, indicating that patients reporting increasing relief from baseline to the end of the treatment period also experienced improvement in the symptoms assessed by the IBS-C Symptom Severity Measures. Correlations between additional PRO measures (IBS-SSS, IBS-QOL, HADS-A and HADS-D) and the IBS-C Symptom Severity Measures demonstrated convergent validity and divergent validity and were similar across these studies. With respect to the rather low correlations between the IBS-C Symptom Severity Measures and the IBS-QOL, it was originally hypothesised that these correlations would be weak. Symptoms and impacts are distinct concepts and it typically takes time before patients report improvement in impacts after improvement in symptoms. The estimated differences in the known group means were in the anticipated direction and statistically significant. Responsiveness effect sizes, based on the linaclotide trials' responder definitions, were strong, suggesting that the IBS-C Symptom Severity Measures are capable of detecting change.

The results of the present psychometric evaluation are consistent with the results of similar psychometric analyses conducted with phase 2b linaclotide trial data[20] and build on the qualitative evidence for the IBS-C Symptom Severity Measures [7] demonstrating that these measures have satisfactory psychometric properties and are reliable, valid and responsive in an IBS-C population. Ensuring the measures used to derive trial endpoints have acceptable psychometric properties supports the conclusion of a treatment benefit for linaclotide in the primary bowel symptom and abdominal symptom endpoints and all key secondary endpoints [8–10] in the clinical studies. While the full set of IBS-C symptoms addressed by the IBS-C Symptom Severity Measures has been identified as important to patients through extensive qualitative research and content validation,[7] currently only two symptoms – abdominal pain and BM frequency – are recommended by the FDA to define primary endpoints in IBS-C trials to support drug approval.[13] However, because the FDA has indicated a preference for the development and use of a multi-item PRO instrument to capture all clinically important symptoms of IBS-C within the primary endpoint of future trials, the factor analyses explored the feasibility of one or more multi-item IBS-C composites. In this study, two highly correlated abdominal symptom and bowel symptom components were observed to optimally explain the pattern of correlations among the set of IBS-C Symptom Severity Measures, suggesting that the development of abdominal and bowel symptom composite scores might be a reasonable alternative to using two single symptoms for defining improvement. In addition, two separate abdominal and bowel symptom composites offer clinical

utility by allowing the detection of a differential treatment impact of different therapies, vital information for physicians, patients and other stakeholders. Future studies should focus on further investigating the feasibility of separate bowel and abdominal symptoms composite scores and constructing optimal scoring algorithms, as well as demonstrating reliability, validity and responsiveness of these measures.

AUTHORSHIP

Guarantor of the article: Valerie Williams, PhD.

Author contributions: Valerie Williams, PhD collaborated on the development of the analysis plan, conducted the analyses, and drafted the initial manuscript, incorporated feedback from co-authors, and prepared the submitted version of the manuscript. Lauren Nelson, PhD collaborated on the development of the analysis plan, conducted the analyses, contributed to the development of the manuscript and responded to feedback from co-authors. Sheri E. Fehnel, PhD contributed to the development of the IBS-C Symptom Severity Measures and collaborated on the development of the analysis plan, reviewed and contributed to the development of the manuscript. Mollie Baird, MPH collaborated with RTI Health Solutions on the development of the analysis plan and the interpretation of results, reviewed and contributed to the development of the manuscript. Robyn Carson, MPH contributed to the development of the IBS-C Symptom Severity Measures and collaborated with RTI Health Solutions on the development of the analysis plan and the interpretation of results, reviewed and contributed to the development of the manuscript. Jeffrey M. Johnston, MD contributed to the development of the IBS-C Symptom Severity Measures and collaborated with RTI Health Solutions on the development of the analysis plan and the interpretation of results, reviewed and contributed to the development of the manuscript. James MacDougall, PhD contributed to the development of the IBS-C Symptom Severity Measures and collaborated with RTI Health Solutions on the development of the analysis plan and the interpretation of results, reviewed and contributed to the development of the manuscript. Stavros Tourkodimitris, PhD collaborated with RTI Health Solutions on the development of the analysis plan and the interpretation of results, reviewed and contributed to the development of the manuscript. Caroline Kurtz, PhD contributed to the development of the IBS-C Symptom Severity Measures and collaborated with RTI Health Solutions on the development of the analysis plan and the interpretation of results, reviewed and contributed to the development of the manuscript. All authors approved the final version of the manuscript.

ACKNOWLEDGEMENTS

Declaration of personal interests: Caroline Kurtz and Jeffrey Johnston are employees of Ironwood Pharmaceuticals, Inc. and own stock/stock options in Ironwood Pharmaceuticals, Inc. James MacDougall and Mollie Baird were employees of Ironwood Pharmaceuticals, Inc. while these analyses were conducted. Robyn Carson and Stavros Tourkodimitris are employees of Forest Research Institute and own stock/stock options in Forest Laboratories, Inc. Valerie Williams, Lauren Nelson and Sheri Fehnel are paid consultants to Ironwood Pharmaceuticals, Inc. and Forest Research Institute. The authors thank Cheryl Coon for her assistance with the factor analysis.

Declaration of funding interests: The study was conducted by RTI Health Solutions, which received consultancy fees from Forest Research Institute and Ironwood Pharmaceuticals, Inc. Forest Research Institute, Ironwood Pharmaceuticals, Inc., and RTI Health Solutions were involved in the analysis and interpretation of data and the decision to submit these data for publication. Funds for preparation of this manuscript were provided by Ironwood Pharmaceuticals, Inc. and Forest Research Institute. RTI Health Solutions is a research unit of RTI International, a not-for-profit research institute.

REFERENCES

1. Longstreth GF, Thompson WG, Chey WD, *et al.* Functional bowel disorders. *Gastroenterology* 2006; **130**: 1480–91.
2. Sandler RS. Epidemiology of irritable bowel syndrome in the United States. *Gastroenterology* 1990; **99**: 409–15.
3. Andrews EB, Eaton SC, Hollis KA, *et al.* Prevalence and demographics of irritable bowel syndrome: results from a large Web-based survey. *Aliment Pharmacol Ther* 2005; **22**: 935–42.
4. Hungin AP, Chang L, Locke GR, *et al.* Irritable bowel syndrome in the United States: prevalence, symptom patterns and impact. *Aliment Pharmacol Ther* 2005; **21**: 1365–75.
5. Lovell RM, Ford AC. Global prevalence of, and risk factors for, irritable bowel syndrome: a meta-analysis. *Clin Gastroenterol Hepatol* 2012; **10**: 712–21.

6. US Department of Health and Human Services. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. December 2009. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf. Accessed March 24, 2010.
7. Fehnel SE, Ervin CM, Lewis BE, Carson RT, Johnston JM. Identifying endpoints for irritable bowel syndrome (IBS) clinical trials: incorporating the patient's voice. Poster presented at the 13th Annual European Congress of the International Society for Pharmacoeconomics and Outcomes Research (ISPOR); November 2010. Prague, Czech Republic.
8. Johnston JM, Kurtz CB, MacDougall JE, *et al.* Linaclotide improves abdominal pain and bowel habits in a phase IIb study of patients with irritable bowel syndrome with constipation. *Gastroenterology* 2010; **139**: 1877–86.
9. Chey WD, Lembo AJ, Lavins BJ, *et al.* Linaclotide for irritable bowel syndrome with constipation: a 26-week, randomized, double-blind, placebo-controlled trial to evaluate efficacy and safety. *Am J Gastroenterol* 2012; **107**: 1702–12.
10. Rao S, Lembo AJ, Shiff SJ, *et al.* A 12-week, randomized, controlled trial with a 4-week randomized withdrawal period to evaluate the efficacy and safety of linaclotide in irritable bowel syndrome with constipation. *Am J Gastroenterol* 2012; **107**: 1714–24.
11. Drossman DA, Corzziari E, Talley NJ, *et al. Rome II: The Functional Gastrointestinal Disorders*, 2nd ed. Degnon Associates: McLean VA, 2000.
12. Heaton KW, O'Donnell LJD. An office guide to whole-gut transit time: patient's recollection of their stool form. *J Clin Gastroenterol* 1994; **19**: 28–30.
13. US Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER). Guidance for industry: irritable bowel syndrome – clinical evaluation of drugs for treatment. 2012. Available at: http://www.fda.gov/downloads/Drugs/Guidances/UCM205269.pdf. Accessed April 30, 2013.
14. Patrick DL, Drossman DA, Frederick IO, *et al.* Quality of life in persons with irritable bowel syndrome: development and validation of a new measure. *Dig Dis Sci* 1998; **43**: 400–11.
15. Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progress. *Aliment Pharmacol Ther* 1997; **11**: 395–402.
16. Snaith RP, Zigmond AS. *The Hospital Anxiety and Depression Scale Manual.* Windsor, Berkshire, England: Nfer-Nelson, 1994.
17. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand* 1983; **67**: 361–70.
18. SAS Institute, Inc. *SAS Proprietary Software*, version 9.2. Cary, NC: SAS Institute, Inc., 2008.
19. Muthén LK, Muthén BO. *Mplus User's Guide*, 6th ed. Los Angeles: Muthén & Muthén, 1998-2010.
20. Carson RT, Nelson LM, Williams V, MacDougall JE, Johnston JM. Psychometric evaluation of patient-reported outcome measures for assessing IBS–C symptom severity and global change: results from a phase 2b study of linaclotide (poster #681). Poster presented at the American College of Gastroenterology 2009 Scientific Meeting; October 2009. San Diego CA.
21. Nunnally JC, Bernstein IH. *Psychometric Theory*, 3rd ed. New York: McGraw-Hill, 1994.
22. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
23. Guyatt GH, Walter SD, Norman G. Measuring change over time: assessing the usefulness of evaluative instruments. *J Chronic Dis* 1987; **40**: 171–8.