Published in final edited form as:
Clin Gastroenterol Hepatol. 2011 November ; 9(11): 957–964.e1. doi:10.1016/j.cgh.2011.06.014.

Patient-Reported Outcomes for Irritable Bowel Syndrome Are Associated with Patients' Ratings of Symptoms and Non-Clinical Factors

Jeffrey Lackner, Psy.D[1], **James Jaccard, Ph.D**[2], **Charles Baum, M.D**[3], **Amanda Smith, M.S**[1], **Tatayna Raby, B.S**[1], **Susan Krasner, PhD**[4], **Leonard Katz, M.D**[1], **Rebecca Firth, MHA**[1], and **Cathrine Powell, M.A**[1]

[1] Division of Gastroenterology, Department of Medicine, University at Buffalo, SUNY, Buffalo, NY, USA

[2] Department of Psychology, Florida International University, Miami, FL

[3] Takeda Pharmaceuticals, Deerfield, IL

[4] Department of Anesthesiology, University at Buffalo, SUNY, Buffalo, NY

Abstract

Background & Aims—Patient-reported outcomes (PROs) are used to gauge the benefit of treatments for functional gastrointestinal disorders, including irritable bowel syndrome (IBS). Commonly used endpoints derived from scales of symptom severity differ in their structure, format, and the extent to which they are based on established, psychometric fundamentals. We evaluated the overlap between 2 measures of IBS symptom severity, documented their association with different symptoms (pain, bloating, defecation), and identified psychological factors that might bias PRO ratings, by affecting how patients interpret IBS symptom severity.

Methods—Ninety-eight patients diagnosed with IBS, based on Rome III criteria, completed the multi-component IBS symptom severity scale and the single-item, UCLA symptom severity scale. Data were collected on pain, bloating, and bowel habits, as well as somatization, sensitivity to arousal symptoms (anxiety sensitivity), and a negative thinking style called pain catastrophizing.

Results—The 2 global scales were correlated with one another (r=0.56); each scale was most strongly associated with variation in abdominal pain. Data were consistent with a model in which pain catastrophizing and somatization influenced 1 or more of patients' judgments of pain, bloating, and/or bowel habits, which then affected the PROs.

Conclusions—Depending on their structure and format, PROs can have different levels of sensitivity to core IBS symptoms and be influenced by psychological and somatic complaints that

Corresponding Author: Jeffrey M. Lackner, Psy.D.Department of MedicineUniversity at Buffalo School of Medicine, SUNYECMC, 462 Grider Street, Buffalo, NY 14215716.898.5671 (tel); 716.898.3040 (fax) lackner@buffalo.edu.

Competing Interests: None

Study concept and design: Lackner, JaccardAcquisition of data: Firth, Krasner, Katz, Lackner, PowellAnalysis and interpretation of data: Lackner, Jaccard Drafting of the manuscript: Lackner, Jaccard, Baum, Smith, Raby Critical revision of the manuscript for important intellectual content: Lackner, Jaccard, BaumStatistical analysis: Jaccard Obtained funding: LacknerAdministrative, technical, or material support: Firth, Smith, Raby, Lackner, Krasner, Katz, PowellStudy supervision: Lackner

are beyond the aim of therapy and labeling claim. PROs that rely on patients' perspectives to index symptom severity can be improved by consideration of psychometric principles that influence self report.

Keywords

outcome assessment; health status indicators; clinical practice; symptom severity

Irritable bowel syndrome (IBS) is a chronic, painful, often times disabling gastrointestinal condition for which there is no satisfactory medical treatment for its full range of symptoms. Two of the past three drug therapies approved by the FDA for the treatment of IBS have required regulatory intervention leading to drug withdrawal in one case and a severely restrictive risk management program in the other. These events, coupled with recent FDA restrictions on the primary study endpoint to be used in IBS pharmaceutical development, have reduced perceived commercial value of new drug development for IBS and limited options for one of the most common gastrointestinal disorders[1].

To meet the unmet need for safe, effective, and widely available treatments for IBS, the FDA has issued a PRO (patient-reported outcome) Guidance document[2–3] that identifies limitations of endpoints used to support claims for previously approved IBS drugs and that elaborates some of the fundamental psychometric properties and processes by which well defined, meaningful, and valid PROs are developed. A PRO instrument is used to capture clinically important information regarding the therapeutic benefit of treatment from the patient perspective. The FDA requirements herald a new era in gastrointestinal drug development for the spectrum of symptom-based gastrointestinal diseases (e.g., functional GI disorders) whose expression is based on self report data. The assessment of drug efficacy using PRO endpoints requires instrument development and validation incorporating psychometric principles that are likely foreign to many within the gastroenterology community. The current work addresses two expert derived instruments that supposedly measure therapeutic benefit in terms of the severity of IBS symptoms using patient rating scales.

One strategy for measuring IBS symptom severity is to have patients respond to a single global severity item in which patients rate the overall severity of their symptoms, an example of which is the UCLA Symptom Severity Scale [4]. In this strategy, the specific clinical domains of interest are left unspecified and patients define severity in whatever way comes to mind at the time the rating is made. Such global ratings are quick to administer, easy to score, and presumably easy to interpret. The approach is thought to have a practical advantage for gauging therapeutic benefit of agents for multisymptom syndromes such as IBS whose symptom complexion differs across patients. Global ratings of symptom severity yield a common metric that ostensibly cuts across the heterogeneity of symptoms of patients with different symptom profiles (e.g., predominant bowel habits). The difficulty, however, is that the clinician/researcher does not know what symptom(s) patients use as the basis for their judgments and they thereby struggle to translate patients ratings into a clear, meaningful metric of treatment response. Furthermore, the rationale for global endpoints rest on the unproven assumption that they are sensitive to the full range of symptoms of IBS and not a single dominant symptom (e.g., pain) at the expense of others, a subset of symptoms, extraintestinal symptoms, and/or coexisting psychological processes that influence symptom perception. Indeed, if a patient experiences multiple symptoms of varying severity, it is unclear how the patient combines judgments of severity of the individual symptoms into an overall global judgment of severity (e.g., whether the patient uses a mental average of the different severities, or, alternatively, the highest experienced severity). Some psychological processes that bear on severity ratings may be long standing, stable, ingrained characteristics

that influence the way patients perceive somatic sensations independent of life circumstances (e.g. undergoing treatment as part of a clinical trial). Finally, the processes underlying the impact of a treatment on one type of symptom may be different than the processes affecting another type of symptom, making it difficult to use such general measures to build a cohesive body of knowledge about symptoms and their responsiveness to novel agents.

Another difficulty with many measures of IBS symptom severity is that they allow patients to define "severity" in their own idiosyncratic ways. Severity judgments are based on how patients define severity, and how they recall symptoms and retrieve information from memory. The recalled symptom information that forms the basis of a severity judgment may be based on any number of factors such as the emotional unpleasantness of symptoms, their impact on quality of life, memories, meaning (e.g., uncontrollability, unpredictability), or perceptual salience[5].

A second assessment strategy for IBS symptom severity is to use composite measures that aggregate patient ratings of different, discretely defined symptom dimensions into a single score. An example is the IBS Symptom Severity Scale (IBS-SSS)[6] which Whitehead and associates describe as "the most comprehensive of the IBS symptom severity scales currently available"[7]. Such "integrative" [8] scales explicitly define the symptoms patients are to consider and then impose an (arbitrary) aggregation algorithm to define how those symptoms combine into an overall global index. Whereas in the former strategy, specification of symptoms and the aggregation algorithm for them is left to the (unknown) mental operations of the patient, the composite approach places these matters more under the control of the researcher. It is unknown if the two approaches generate comparable characterizations of symptom severity. A purpose of this study is to examine the convergence between them. In addition, we evaluated the extent to which each scale is sensitive to variation in three types of GI symptoms, (1) pain, (2) bloating, and (3) defecatory symptoms.

Obtaining accurate reports about symptoms is influenced by multiple factors. These include qualites of the examiner, the respondent, and the task itself. Respondent factors include his or her psychological makeup. We examined the extent to which such factors in IBS patients are associated with PRO endpoints of severity. Specifically, we investigated three constructs associated with the experience of having IBS: pain catastrophizing, somatization, and anxiety sensitivity. Pain catastrophizing is the tendency to focus on and exaggerate pain experience either in direct response to pain or in anticipation of painful stimuli [9–11]. To the extent that pain catastrophizing impacts subjective experiences of pain and to the extent that subjective experiences of pain are a basis for patient judgments of general symptom severity, then classic PROs for symptom severity should reflect the influence of pain catastrophizing. Somatization refers to unexplained bodily complaints [12] (e.g. dizziness, benign headache, back pain, fatigue) involving multiple organ systems and is influenced by psychological factors, such as stress reactivity, illness beliefs, and negative emotionality. Patients who somatize report unexplained physical symptoms of which IBS symptoms is common [13]. As such, somatization also could impact IBS symptom ratings. In other words, unexplained medical complaints of an extraintestinal nature stand to influence the severity of GI symptom ratings used to measure IBS targeted therapies. Anxiety sensitivity is a relatively stable personality trait marked by a fear of somatic symptoms arising from the belief that they have harmful consequences [14]. Descriptively similar to the construct of visceral sensitivity [15], anxiety sensitivity (AS) involves heightened vigilance to arousal symptoms as well as catastrophic interpretations of the meaning of somatic symptoms. Individuals high in AS tend to selectively attend to physically threatening stimuli and are more hypervigilant to noxious bodily symptoms [14]. To the extent that IBS symptom severity

ratings tap these symptoms, then high AS individuals would be expected to report more severe levels of IBS symptoms.

Method

Participants

Participants included 98 consecutively evaluated Rome II-diagnosed IBS[16] patients recruited at an academic medical center. Subject characteristics are presented in Table 1.

Procedure

After a brief telephone interview to determine eligibility, participants completed psychological testing, which included the measures described below. Detailed information about study procedure can be found elsewhere[17]

Measures

Irritable Bowel Syndrome Symptom Severity Scale (IBS-SSS)—The IBS-SSS[18] is a 5-item global measure of IBS symptom severity based on descriptions of abdominal pain, frequency of abdominal pain, severity of abdominal distension, dissatisfaction with bowel habits, and interference with quality of life. Each item uses a 100-point scale. For four items, respondents mark a point on the line reflecting the extremity of their judgment and the proportional distance from zero is the score assigned for that item (hence scores range from 0 to 100). A final item asks the number of days out of 10 the patient experiences abdominal pain with the answer multiplied by 10 to create a 0 to 100 metric. The five items are summed to yield a total score of 0 to 500.

UCLA Global Severity of Gastrointestinal Symptoms Scale (UCLA SS)—The UCLA SS[4] is a single item 21-point rating scale that measures severity of IBS symptoms. Participants rate symptoms from 0 (no symptoms) to 20 (most intense symptoms imaginable).

UCLA Pain Scale—Patient experience of pain was measured using the 21-point UCLA scale that measures pain severity[4] . Participants rated their abdominal pain severity during the past week from 0 (no pain) to 20 (the most intense pain imaginable).

Altered defecation—Defecatory symptoms were measured using the non pain bowel symptom items of the Gastrointestinal Symptoms Rating Scale –IBS (GSRS-IBS;[19]). The GSRS- IBS is a 13-item measure that assesses severity of different gastrointestinal symptoms associated with IBS. Each item is evaluated on a seven-point scale (1 = no discomfort, 7 = very severe discomfort). For this study, two distinct factors were created, a bloating/distention factor that averaged two GSRS items and a defecation factor that averaged defecation items (ease of stool passage, stool consistency, urgent need for defecation, feeling of incomplete evacuation).

Pain Catastrophizing—The six items of the Catastrophizing subscale of the CSQ [20] asks patients to rate the frequency with which they engage in thoughts that index catastrophizing during pain episodes (e.g., “When I am in pain, I feel I can’t stand it anymore),. Respondents rate each item using a scale ranging from 0 (never do) to 7 (always do).

Anxiety Sensitivity Inventory—The ASI [21] is a self-report measure that reflects fear of anxiety (e.g., “It scares me when I am anxious”), arousal related bodily sensations (“It scares me when my heart beats rapidly”), and their consequences (e.g., “When I notice my heart is

beating rapidly, I worry that I might have a heart attack"). Each of the 16 items of the ASI is rated on a six point scale (0 =very little, 5 = very much).

Somatization—Somatization was measured using the seven-item Somatization Scale of the Brief Symptom Inventory[22–23]. The Somatization Scale measures distress arising from perceptions of bodily dysfunction in multiple physical systems (e.g., cardiovascular, respiratory, musculoskeletal). Items are rated on a five point scale (0 – not at all, 4 extremely) to reflect somatic distress during the previous week.

Data Analysis

Analyses were guided by the model using the traditional path diagram in Figure 1. Variables are represented by boxes and presumed causal relationships by arrows. The box an arrow emanates from is the presumed cause and the box the arrow points to is the presumed effect. Presumed influences can be direct (where one variable is directly connected to another variable) or indirect (where a variable is thought to influence another variable through an intermediary variable). The endpoint of symptom severity (either the UCLA or IBS-SSS measure) was conceptualized as a direct linear function of three types of GI-associated symptoms, pain, bloating, and defecatory symptoms. Each of these symptoms, in turn, was conceptualized as being a direct linear function of pain catastrophizing, somatization, and anxiety sensitivity. We tested an alternative model that posited that the psychological variables directly impacted overall severity scores independent of their influence on symptom interpretation but found no support for it (i.e., we observed non-significant path coefficients for the direct effects), so it is not considered further. Of primary interest are the regression coefficients for the linear equations implied by Figure 1. These coefficients were estimated using multiple strategies, as detailed in the Appendix. In addition to unstandardized regression coefficients, we report statistics that provide perspectives on the relative contributions of each predictor to a PRO. These include zero order correlations, semi-partial correlations, standardized regression coefficients, and indices of relative importance using methods suggested by Johnson[24–25] (see Appendix). Details surrounding missing data, outliers, non-normality, and covariates are in the Appendix.

To place variables on a common metric and to facilitate interpretation of unstandardized coefficients, all variables in Figure 1 were converted to a 0 to 10 metric (i.e. we transformed means and standard deviations but without altering the distributional properties of the measures). We subtracted the lowest possible score attainable for each variable from an individual's observed score and then divided this by the highest possible score on that variable. This result was multiplied by 10, yielding scores from 0 to 10, with 5 representing a scale midpoint.

Results

GI Symptoms and the Global Measure of Symptom Severity

The leftmost columns of Table 2 present the unstandardized coefficients and the indices of predictor importance when predicting the general rating of symptom severity based on the UCLA patient reported outcome from the three types of GI symptoms. Analyses also were conducted separately for subgroups representing each of the three major bowel subtypes of IBS (IBS-C, IBS-D, IBS-mixed). The trend in these subgroup analyses mirrored the magnitude of coefficients for the total sample, hence only the latter are reported. The three IBS symptom variables, taken together, yielded a multiple correlation of 0.71 for predicting the PRO measure, accounting for about 50% of the variation. This suggests a large proportion of the variation in the overall symptom severity scores can be accounted for by these symptoms, but that there also is a large proportion of variation that is not attributable

to them. The IBS symptom of abdominal pain was the dominant predictor of the UCLA rating of symptom severity. In general, for every one unit that the rating of pain increased on its 0 to 10 metric, the predicted mean of the overall severity of symptoms increased by 0.46 units on it 0 to 10 metric, or about half a scale unit ($t(97) = 6.46$, $p < 0.05$). For the defecatory symptoms of IBS, for every one unit that the rating of defecatory symptoms increased on its 0 to 10 metric, the predicted mean of the overall severity of symptoms increased by 0.23 units on its 0 to 10 metric, or about one quarter of a scale unit ($t(97) = 2.20$, $p < 0.05$). The coefficient for bloating was not statistically significant.

GI Symptoms and the Multi-Item Measure of Symptom Severity

The rightmost columns of Table 2 present the unstandardized coefficients and various indices of predictor importance when predicting the IBS-SS scale from the three types of GI symptoms. As with the UCLA measure, the trends in the IBS subgroups were similar to the total sample. Taken together, the three IBS symptom variables yielded a multiple correlation of 0.76 when predicting the IBS-SS measure and accounted for 57% of the variation. Thus, about 43% of the variation in the scores is due to unspecified factors not associated with the symptoms. Abdominal pain was again the dominant predictor of the overall severity score. For every one unit that the rating of pain increased on its 0 to 10 metric, the predicted mean of the overall severity of symptoms increased by 0.34 units on it 0 to 10 metric, or about one third a scale unit ($t(97) = 6.88$, $p < 0.05$). For bloating, for every one unit that the rating of bloating increased on its 0 to 10 metric, the predicted mean of the overall severity of symptoms increased by 0.14 units on its 0 to 10 metric ($t(97) = 3.41$, $p < 0.05$). The coefficient associated with defecatory symptoms was not statistically significant.

Convergence between Scales

The correlation between the two IBS symptom severity measures was 0.59 ($p < 0.05$). Thus, one measure accounts for about 36% of the variation in the other measure. Stated another way, if we completely remove variation in the IBS-SS that can be accounted for by the UCLA measure, or vice versa, about 64% of the original variation in the scale that remains is due to other factors. The mean absolute discrepancy between the UCLA severity measure and the IBS-SSS measure was 1.36 scale units on a 0 to 10 metric, and the median absolute discrepancy was 1.13, indicating that the disagreement between the two measures was, on average, about 1 full scale unit on the 0 to 10 metric. About 21% of respondents had a discrepancy of ± 2.0 scale units on the 0 to 10 metric. Figure 2 presents side-by-side histograms for each severity outcome on their 0 to 10 metric and a superimposed normal distribution on each histogram.

The Contribution of Psychological Factors to PROs

Table 3 presents the results that regress each of the GI symptoms onto the three psychological predictors of pain catastrophizing, somatization, and anxiety sensitivity. Noteworthy in these analyses is the consistent role that somatization has in predicting each GI symptom, with higher scores on somatization associated with higher ratings of IBS symptoms. Pain catastrophizing also was a statistically significant predictor of GI symptoms related to pain ($p < 0.05$).

The model in Figure 1 assumes that the psychological factors impact both the overall UCLA and the IBS-SSS endpoints, but that they do so indirectly through their effects on patient experiences of the three GI symptoms. The presumed overall impact of a given psychological factor on a global PRO can be estimated by multiplying the unstandardized regression coefficients in the mediational chain and summing these across the different chains. For example, combining Tables 2 and 3, the estimated impact of somatization on the UCLA global severity rating is (0.46)(0.61) + (0.07)(0.66) + (0.23)(0.25) = 0.38. Thus, for

every one unit that somatization increases on its 0 to 10 metric, the UCLA global severity rating is predicted to increase, on average, by .38 units on its 0 to 10 metric, or just over 1/3 of a scale unit. For the IBS-SS global severity rating, the overall estimated effect of somatization on it is 0.33.

Discussion

The present study explored the psychometric properties of two PRO scales used in the IBS literature, one a single-item global rating of IBS symptom severity and the other a composite multi-item scale that explicitly incorporates different experiences of IBS. Several interesting results emerged. First, a substantial proportion of the variation in the PROs (50% to 55%) could be explained by three distinct GI symptoms of pain, bloating, and defecation symptoms. Despite this, considerable variability on the scales (about 45% to 50%) was due to factors other than these three symptoms, at least as operationalized and estimated in this research. The symptom of abdominal pain was the dominant predictor of both endpoints, with each scale showing somewhat different sensitivity to bloating versus defecatory symptoms (though larger sample sizes and additional research are required to affirm this). Neither endpoint seems to be impacted as much by defecatory symptoms which are, of course, core features of IBS, troublesome for patients, and targets of both pharmacological and behavioral therapies. Another interesting finding concerned the degree of convergence between the data of single item global rating scale (UCLA) and the averaged multi-item IBS-SSS scale. The correlation between them was 0.59. Although the measures share substantial common variance, there also is a non-trivial amount of unique variance in them. The measures tap into common but also somewhat different phenomena. The histograms in Figure 1 suggest that even though they ostensibly tap the same construct (IBS symptom severity), the measures have somewhat distributional differences, especially at the lower end of the scales, where there are higher frequencies for the UCLA as opposed to the IBS-SS scale (in supplemental analyses we performed, we found statistically significant disparities in the lower deciles of the two scales). Whether such disparities matter at a clinical level is unknown because the metrics of both scales are, at this stage of research, arbitrary.

Traditionally, the quality of PROs has been documented in terms of their reliability and validity[26]. Although important, reliability and validity are not the sole necessary criteria for a measure to be meaningfully interpretable[27]. The metric of the measure also must be non-arbitrary. According to Blanton and Jaccard[28–29], an arbitrary metric is one whose numerical values are not clearly and concretely linked to specific locations on the underlying dimension being measured. Arbitrary metrics cannot, for example, indicate how much change – the ultimate goal of clinical research -- has been produced *on the underlying construct* (e.g. symptom severity, quality of life) given an observed numerical change on the scale itself. A classic example of metric arbitrariness is height as measured in meters for the typical American citizen. Meters is a well defined, completely reliable and valid index of height and, in this sense, using meters to describe height for American citizens would clear the psychometric bar the FDA sets for PROs. The problem is that most Americans have no clue about how tall someone is who is 1.90 meters tall nor how much taller the individual is as compared to someone who is 1.77 meters tall. It is only when this numerical system is tied to concrete, meaningful benchmarks on the underlying dimension of height that the meaning of the scale numbers become non-arbitrary and interpretable. The same problem extends to the concept of symptom severity and the numerical values severity scales yield. What does a numerical rating of 5 mean on the 0 to 20 metric of the UCLA scale? And what does it mean to shift someone from a score of 9 to a score of 4 on that scale? What does a score of 210 on the IBS-SSS imply about someone's severity of GI symptoms and how does this differ from the symptom severity implied by a score of 280? These scores -- much like the metric system is for most Americans -- are arbitrary and of limited value in informing

stakeholders (NIH, regulatory agencies, health care policy makers, patients, providers) about the nature and extent of treatment effects. The issue of metric arbitrariness is particularly relevant for clinicians whose skepticism about the clinical meaning of PRO assessment data prevents their integration into daily clinical practice [30–31].

We are not convinced that applying data transformations to index clinical significance (e.g., deriving standardized effect size estimates or MCIDs[32], calibrating the observed score on a scale to values for some reference group[33], examining scores in terms of standard deviation units[34]) render a scale value any less arbitrary and any more meaningful than the raw metrics (for elaboration see[35]). Reporting normative data on subjective constructs such as symptom severity or quality of life specifies where an individual stands on a given metric relative to other people (e.g., MCID of the IBS-QOL ≥ 14, half SD improvement, 50 point pre-post treatment reduction on IBS-SS) but this information is not inherently informative about whether the derived value amounts to a real or important change brought on by a treatment[35]. By the same token, quantitative indices used in IBS trials such as number of bowel movements or stool type , while important in their own right, take on an arbitrary quality when they are used to infer psychologically elaborated constructs such as magnitude of symptom severity, therapeutic benefit or relief, quality of life, or satisfaction. Clearly there is reason to look beyond the two classic properties of reliability and validity as psychometric hallmarks for PRO measures. We suggest the addition of a third hallmark, namely metric meaningfulness. Future research that calibrates specific numbers of symptom scales to tangible, meaningful benchmarks is necessary to advance the field so that we can approach our metrics with the same comfort and confidence that we approach such metrics as dollars, inches, and pounds.

The present study also explored the potential contribution of psychological factors to patient reported experiences of GI symptoms, specifically the tendency to catastrophize, somatize, and the tendency to misinterpret noxious bodily sensations (i.e. anxiety sensitivity). A robust finding was the role of somatization (unexplained boldly complaints involving multiple organ systems) as a correlate of patient reported experiences surrounding the IBS symptoms of pain, bloating, and altered defecation. Although somatization was not directly associated with the overall PROs independent of these symptoms, the data are consistent with the view that it still may affect the overall PROs indirectly by shaping the kinds of subjective experiences patients reported for pain, bloating, and defecatory symptoms. This suggests that existing PROs may advantage behavioral interventions because their techniques[17, 35] are better able to contextualize the full range of health problems in a way that impacts overall symptom severity ratings. This result also highlights the value of biobehavioral treatments that address not only the physical symptoms of IBS per se, but the way that people experience, interpret, and respond to their symptoms. Because IBS PROs apply to treatments with underlying psychological and physical mechanisms, it may sometimes be desirable to develop more fine-grained PROs that are differentially sensitive to the multiple mechanisms. For example, when evaluating medications to alleviate bowel symptoms, it is reasonable to expect that the effects of the medications exert a more proximal effect on the physical symptoms per se rather than the cognitive processes such as patient interpretations of those symptoms. This proposition echoes the FDA's focus[26] on developing PROs that are proximally linked to a disease's cardinal symptoms for which a labeling claim is being sought. Our results suggest that levels of distress about extraintestinal symptoms (i.e. somatization) might color severity ratings of IBS symptoms used to characterize the therapeutic benefit of agents explicitly designed for intestinal symptoms of IBS. If extraintestinal symptoms influence patient ratings of IBS symptoms, there is reason to consider the value of "purifying" PROs by, for example, obtaining measures of such constructs and then applying statistical methods for covarying out the contaminating influence of somatization and other psychological factors. At the very least, our data indicate

that existing PROs pick up, to some extent, extraintestinal somatic complaints and possibly stable, deeply ingrained personality traits that are not the intended focus of labeling claims for agents the PROs seeks to assess. An innovative implication of our findings is the potential value of assessing the extent to which a broad range of biobehavioral factors systematically impact treatment response early in drug development (Phase II) and then to use these data to develop integrated treatments that tackle the full range of underlying disease mechanisms (e.g. central and peripheral) beyond the scope of any one treatment per se.

Our results invite consideration of factors that may influence symptom severity ratings of PROs beyond the patient factors featured in the present study. One important and largely overlooked set of factors that influence self report data relates to the features of the instruments used to capture self report PRO data (eg item wording, content, questionnaire context). Cognitive research suggests that self-report instruments that comprise PROs are hardly a passive, neutral receptacle of patient information. Nor is answering questions a reflexive, straightforward, precise process[36]. Reporting one's symptoms is a cognitively demanding task that involves multiple processes[36–38]. These tasks include *comprehension*, where the respondent interprets the gist (i.e., meaning) of the questions; *retrieval, w*here they recall from long term memory the relevant information; *judgment/estimation*, where they combine what they have retrieved from memory to form a unitary mental judgment; and *response translation*, where they translate the judgment into an answer category provided on the rating scale, perhaps altering it for consistency with prior answers, social desirability, or other criteria.. Each of these processes influence the value patients assign to a scale to describe their symptoms. Two patients may extract the same meaning with respect to a question and form the same judgment, yet provide different numerical ratings if they interpret the rating scales differently. Similarly, patients may interpret the same question differently, leading to different judgments and numerical responses as they respond to what are functionally different questions even thought the words in the question are identical. This means that a patient's response to questions about their symptoms -- whether posed in a clinical or research setting -- may be far from an accurate "snapshot" of internal states (e.g., symptoms) but also may be obscured by "noise" (i.e., measurement error) that emanates at each response process. A seemingly minor change in wording, content, order, or format can impact the quality of data. For this reason, clinical researchers whose work relies on PROs must attend to all four mental tasks to arrive at an accurate answer about patients' symptom status. Overlooking these processes runs the risk of blunting the sensitivity of PROs and their ability to detect treatment effects for novel agents.

The PRO movement is likely in the near future to extend to clinical settings where a premium is placed on understanding symptoms from the patient's perspective. Clinicians who face pressure to deliver evidence based medicine in a cost effective manner within tight time constraints at a level that is satisfying to patients and other stakeholders are likely to find that PRO data can increase their understanding of what patients mean when they describe how they function or feel. Greater integration of PRO data into clinical practice can improve decision-making, track disease progression, identify at risk patients, promote physician- patient relationship, and provide a mechanism for ongoing monitoring of quality of care. For PROs to support the efforts of clinical gastroenterologists, it will be important to recognize the limitations inherent in the conventional approach of asking symptom questions, to understand the reasons for these shortcomings, and to develop concrete, systematic ways to minimize their impact and maximize self report accuracy. This study makes clear that attending to the science of asking and answering questions is an inescapable but potentially positive step in the direction of understanding patients' symptoms and building better PROs that accurately characterize the true benefits of prescriptive treatments.

The results of the present research, of course, must be interpreted in light of limitations with the data and statistical modeling. The sample was small in size and restricted to participants who volunteered for a clinical research study at a single specialty clinic in an academic medical setting in the Northeast. The measures are subject to measurement error, which can bias parameter estimates in the regression analyses. While we used a well-regarded measure of somatization[39], it is, like most somatization measures, unable to parse out the organic vs. functional nature of somatic complaints. Therefore our somatization measure should be interpreted conservatively as a measure of distress associated with multiple somatic symptoms[40] and not severity of medically unexplained symptoms. The analyses were associational in nature thereby limiting causal conclusions. The model in Figure 1 assumed directional causation, but it is plausible that some degree of reverse causation also is operative. Despite such limitations, the results of the research are interesting and set the stage for more refined empirical work on (a) understanding the cognitive processes patients use to generate responses on IBS- symptom severity rating scales, (b) making the metrics of these scales non-arbitrary and more clinically meaningful, (c) addressing the role of stable psychological factors that impact subjective experiences of symptoms and, hence, complicate the evaluation of agents to reduce symptom severity, and (d) how best to aggregate, if at all, indices of symptom severity into an overall index by balancing the mental calculus of the patient with the explicit, evidence-based calculus of medical experts whose work relies on the integrity of patient reported endpoints.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Grant support: This research was supported in part by research grants 67878 and 77738 from the National Institutes of Health/NIDDK.

Abbreviations used

IBS	irritable bowel syndrome
PRO	patient reported outcome
IBS	irritable bowel syndrome
IBS-SSS	Irritable Bowel Syndrome Symptom Severity Scale
UCLA-SSS	UCLA Symptom Severity Scale
SRS	Gastrointestinal Symptoms Rating Scale
CSQ	Coping Strategies Questionnaire
ASI	Anxiety Sensitivity Inventory
MCID	Minimal Clinically Important Difference

References

1. Mayer EA. Clinical practice. Irritable bowel syndrome. N Engl J Med. 2008; 358:1692–9. [PubMed: 18420501]
2. Food and Drug Administration. U.S. Department of Health and Human Services. Guidance for Industry Irritable Bowel Syndrome — Clinical Evaluation of Products for Treatment. 2010.

3. U.S Department of Health and Human Services Food and Drug Administration. Guidance for Industry: Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
4. Spiegel B, Strickland A, Naliboff BD, Mayer EA, Chang L. Predictors of patient-assessed illness severity in irritable bowel syndrome. Am J Gastroenterol. 2008; 103:2536–43. [PubMed: 18637089]
5. Redelmeier DA, Katz J, Kahneman D. Memories of colonoscopy: a randomized trial. Pain. 2003; 104:187–94. [PubMed: 12855328]
6. 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. [PubMed: 9146781]
7. Whitehead WE, Palsson OS, Levy RL, Feld AD, VonKorff M, Turner M. Response to Dr. Allen Mangel's Letter. Am J Gastroenterol. 2006; 101:2885–2886.
8. Irvine EJ, Whitehead WE, Chey WD, Matsueda K, Shaw M, Talley NJ, Veldhuyzen van Zanten SJ. Design of Treatment Trials for Functional Gastrointestinal Disorders. Gastroenterology. 2006; 130:1538–1551. [PubMed: 16678567]
9. Sullivan M, Thorn B, Haythornthwaite JA, Keefe FJ, Martin M, Bradley LA, Lefebvre JC. Theoretical perspectives on the relation between catastrophizing and pain. Clin J Pain. 2001; 17:52–64. [PubMed: 11289089]
10. Lackner JM, Quigley BM. Pain catastrophizing mediates the relationship between worry and pain suffering in patients with irritable bowel syndrome. Behav Res Ther. 2005; 43:943–57. [PubMed: 15896288]
11. Lackner JM, Quigley BM, Blanchard EB. Depression and abdominal pain in IBS patients: The mediating role of catastrophizing. Psychosomatic Medicine. 2004; 66:435–441. [PubMed: 15184708]
12. Whitehead WE, Palsson O, Jones KR. Systematic review of the comorbidity of irritable bowel syndrome with other disorders: what are the causes and implications? Gastroenterology. 2002; 122:1140–56. [PubMed: 11910364]
13. Smith GR Jr, Monson RA, Ray DC. Patients With Multiple Unexplained Symptoms: Their Characteristics, Functional Health, and Health Care Utilization. Arch Intern Med. 1986; 146:69–72. [PubMed: 3942467]
14. Taylor, S. Anxiety Sensitivity: Theory, Research, and Treatment of the Fear of Anxiety. Erlbaum; 1999.
15. Labus JS, Mayer EA, Chang L, Bolus R, Naliboff BD. The central role of gastrointestinal-specific anxiety in irritable bowel syndrome: further validation of the visceral sensitivity index. Psychosom Med. 2007; 69:89–98. [PubMed: 17244851]
16. Drossman, DA.; Corazziari, E.; Talley, NJ.; Thompson, GW.; Whitehead, WE.; Rome, II. The functional gastrointestinal disorders. Diagnosis, pathophysiology and treatment: A multinational consensus. Degnon Associates; 2000.
17. Lackner JM, Jaccard J, Krasner SS, Katz LA, Gudleski GD, Holroyd K. Self-administered cognitive behavior therapy for moderate to severe irritable bowel syndrome: clinical efficacy, tolerability, feasibility. Clin Gastroenterol Hepatol. 2008; 6:899–906. [PubMed: 18524691]
18. 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. [PubMed: 9146781]
19. Svedlund J, Sjodin I, Dotevall G. GSRS--a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci. 1988; 33:129–34. [PubMed: 3123181]
20. Rosenstiel AK, Keefe FJ. The use of coping strategies in chronic low back pain patients: relationship to patient characteristics and current adjustment. Pain. 1983; 17:33–44. [PubMed: 6226916]
21. Peterson, RA.; Reiss, S. Anxiety Sensitivity Index: Revised test manual. IDS Publishing Corporation; 1993.
22. Derogatis, LR. Brief Symptom Inventory. National Computer Systems; 1993.

23. Ciccone DS, Just N, Bandilla EB. Non-organic symptom reporting in patients with chronic non-malignant pain. Pain. 1996; 68:329–41. [PubMed: 9121822]
24. Johnson J. A Heuristic Method for Estimating the Relative Weight of Predictor Variables in Multiple Regression. Multivariate Behavioral Research. 2000; 35:1–19.
25. Johnson J, Lebreton J. History and use of relative importance indices inorganizational research. Organizational Research Methods. 2004; 7:238–257.
26. Trentacosti AM, He R, Burke LB, Griebel D, Kennedy DL. Evolution of clinical trials for irritable bowel syndrome: issues in end points and study design. Am J Gastroenterol. 2010; 105:731–5. [PubMed: 20372121]
27. Lackner JM, Jaccard J, Mangel AW, Naliboff BD. Improving the adequacy of the adequate relief measure of clinical efficacy: A Bayesian and conditional probability analysis of a multicenter RCT for HIS. Gastroenterology. 2008; 134:A547–A548.
28. Blanton H, Jaccard J. Arbitrary metrics in psychology. Am Psychol. 2006; 61:27–41. [PubMed: 16435974]
29. Blanton H, Jaccard J. Arbitrary metrics redeux. American Psychologist. 2006b; 61:62–71.
30. Fitzpatrick R, Fletcher A, Gore S, Jones D, Spiegelhalter D, Cox D. Quality of life measures in health care. I: Applications and issues in assessment. BMJ. 1992; 305:1074–7. [PubMed: 1467690]
31. Valderas JM, Kotzeva A, Espallargues M, Guyatt G, Ferrans CE, Halyard MY, Revicki DA, Symonds T, Parada A, Alonso J. The impact of measuring patient-reported outcomes in clinical practice: a systematic review of the literature. Qual Life Res. 2008; 17:179–93. [PubMed: 18175207]
32. Spiegel, BM.; Camilleri, M. Psychometric Evaluation of Binary Endpoints in IBS Clinical Trials: Proposed Analysis Plan for Rome Foundation Outcomes and Endpoints Working Group. Rome Foundation; 2007.
33. Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB. Methods for defining and determining the clinical significance of treatment effects: description, application, and alternatives. J Consult Clin Psychol. 1999; 67:300–7. [PubMed: 10369050]
34. 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; 41:582–92. [PubMed: 12719681]
35. Kazdin AE. Arbitrary metrics: implications for identifying evidence-based treatments. Am Psychol. 2006; 61:42–9. discussion 62–71. [PubMed: 16435975]
36. Schwarz N. Self-reports: How the questions shape the answers. American Psychologist. 1999; 54:93–105.
37. Eisenhower, D.; Mathiowetz, NA.; Morganstein, D. Recall Error: Sources and Bias Reduction Techniques. In: Biemer, P., editor. Measurement Error in Surveys. New York: Wiley and Sons; 1991.
38. Tourangueau, R. Cognitive Science and Survey Methods. In: Jabine, T.; Straf, M.; Tanur, J.; Tourangeau, R., editors. Cognitive Aspects of Survey Methodology: Building a Bridge between Disciplines. Washington, DC: National Academy Press; 1984.
39. Spiegel BM, Kanwal F, Naliboff B, Mayer E. The impact of somatization on the use of gastrointestinal health-care resources in patients with irritable bowel syndrome. Am J Gastroenterol. 2005; 100:2262–73. [PubMed: 16181379]
40. Whitehead WE, Palsson O, Jones KR. Systemic review of the comorbidity of irritable bowel syndrome with other disorders: What are the causes and implications? Gastroenterology. 2002; 98:1187–1192. [PubMed: 2323511]

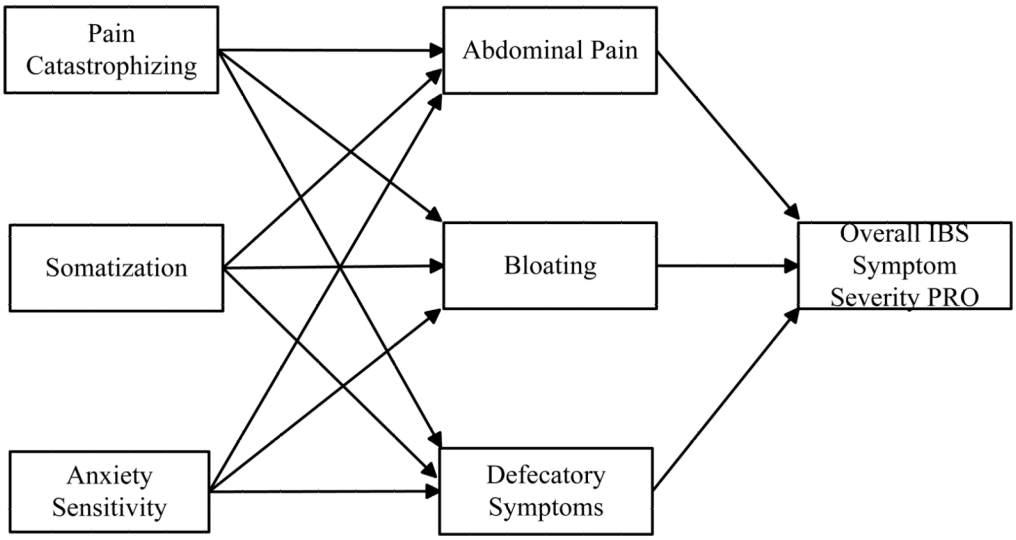

Figure 1.
Path Model Guiding Data Analysis
Note. Psychological factors impact IBS PROs indirectly through their effects on patient experiences of the three GI symptoms

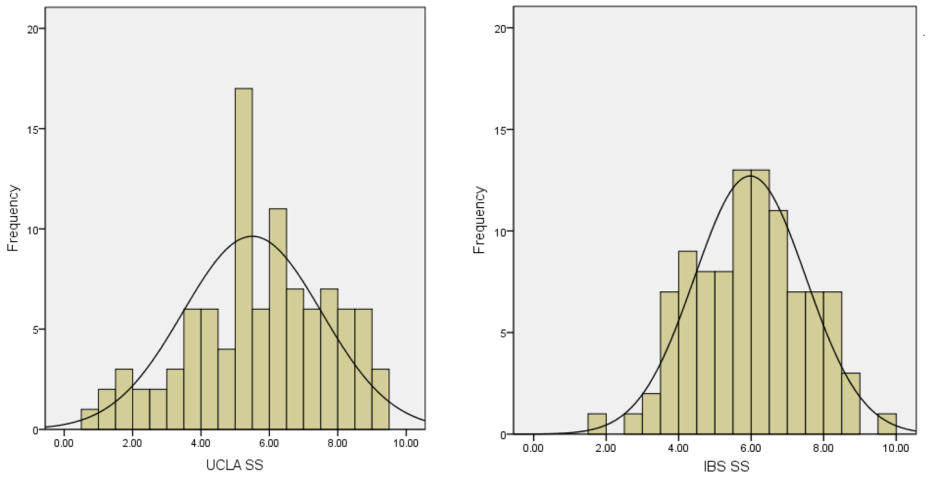

Figure 2.

Side-By-Side Histograms for PRO Outcomes

On their respective 0 to 10 metrics, mean of UCLA = 5.49, SD = 2.03;mean IBS-SS = 5.97, SD 1.54

Table 1

Descriptive Statistics

	Mean	Std. Dev
Percent female	88%	-
Percent married	29%	-
Percent employed full time	46%	-
Percent graduating college	52%	-
Age (yr)	46.7	16.9
Months with IBS	197.8	184.0
Symptom: Pain	4.71	2.48
Symptom: Bloating	4.45	2.79
Symptom: Defecatory symptoms	4.33	1.56
UCLA overall severity	5.49	2.03
IBS-SS overall severity	5.97	1.54
Pain catastrophizing	3.08	2.03
Anxiety sensitivity	3.25	1.59
Somatization	1.89	1.42
IBS Subtypes		
IBS-D 41		
IBS-C 39		
IBS-M 18		
Missing 2		

Note: Untransformed mean for UCLA-SS is 11.52 (transformed mean = 5.49) and for IBS-SSS it is 298.5 (transformed mean = 5.97)..

Clin Gastroenterol Hepatol. Author manuscript; available in PMC 2012 November 1.

Table 2

Results of Regression Analyses Predicting Overall Severity from Three GI Symptoms

	UCLA Overall Severity			IBS-SSS Overall Severity		
	Pain	**Bloating**	**Defecatory Symptoms**	**Pain**	**Bloating**	**Defecatory Symptoms**
Unstandardized coefficient	.46*	.07	.23*	.34*	.14*	.12
Standardized coefficient	.57*	.10	.17*	.55*	.26*	.12
Squared correlation	.46*	.14*	.18*	.50*	.27*	.16*
Squared semi-part correlation	.23*	.01	.03*	.22*	.05*	.01
Relative importance index	67.3%	13.7%	19.0%	60.3%	26.2%	13.5%

Notes:

* $p < 0.05$; significance tests not available for relative importance index

Table 3

Regression Analyses Predicting GI Symptoms from Psychological Variables

	Pain Symptoms			Bloating Symptoms			Defecatory Symptoms		
	Catast	Somat	AS	Catast	Somat	AS	Catast	Somat	AS
Unstandardized coefficient	.27*	.61*	−.11	.09	.66*	−.12	−.16	.25*	.19
Standardized coefficient	.22*	.35*	−.07	.06	.33	−.07	−.21	.23*	.20
Squared correlation	.08*	.15*	.02	.02	.11*	.005	.003	.06*	.04
Squared semi-part correlation	.04*	.10*	<.01	.003	.10*	.003	.03	.04*	.03
Relative importance index	31%	65%	4%	8%	89%	3%	17%	50%	33%

Notes: Catast = Pain catastrophizing, Somat = Somatization, AS = Anxiety Sensitivity;

* $p < 0.05$; significance tests not available for relative importance index: Squared multiple correlations are 0.19 for pain symptoms, 0.11 for bloating symptoms, and 0.11 for defecatory symptoms.