



EDITORIAL

Polarizing or paralyzing? Moving forward with patient reported outcome measurement in irritable bowel syndrome[☆]

¿Polarizar o paralizar? Moviéndonos hacia variables de desenlace reportadas por pacientes en síndrome de intestino irritable

Patients typically seek health care because they experience symptoms. Health care providers must elicit, measure, and interpret patient symptoms as part of their clinical evaluation. Patient-generated reports, also known as Patient-Reported Outcomes (PROs), capture the patients' illness experience in a structured format and may help bridge the gap between patients and providers. The United States Food and Drug Administration (FDA) defines a PRO as "any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else."¹

PROs measure aspects of patient-reported health (e.g. physical, emotional, or social symptoms) and can help to direct care and improve clinical results. When clinicians systematically collect patient-reported data in the right place at the right time, PRO measurement can effectively aid in detection and management of conditions,^{2,3} improve satisfaction with care,⁴ and enhance the patient-provider relationship.⁴⁻⁸

In addition to their use in clinical practice, PROs also play an important role in clinical trials and other research endeavors. For example, health related quality of life (HRQOL), a sub-type of PRO that measures biopsychosocial health, has gained traction as an outcome in clinical research, including clinical trials. HRQOL measures can document patient improvement or decrement over time, and help to estimate the benefits of clinical interventions. In addition, the FDA now considers the patient report in drug approval, and has developed guidance for use of PROs in clinical trials.¹ The National Institute of Health (NIH) has also supported a major PRO initiative, called the Patient Reported Outcome Measurement Information System

(PROMIS[®]; www.nihpromis.org), designed to develop and evaluate several PRO domains.^{9,10} Our group is developing the GI symptom measures within PROMIS. Finally, the rising prominence of the Chronic Care Model, which emphasizes the centrality of the provider-patient relationship in clinical decision making,^{2,3} and places the patient report in the forefront of activity or consideration of health care. In short, there is a confluence of scientific, regulatory, and political factors that amplify the importance of PRO research.

Gastrointestinal (GI) illnesses can lead to physical, mental, and social distress.¹¹ For this reason, patients, providers, investigators, and regulators are interested in using PROs to guide clinical decision-making, conduct clinical research, and achieve drug approval in GI. The conceptual framework in Fig. 1 represents our current understanding of GI symptoms, based on research previously performed by our group in patients with irritable bowel syndrome (IBS).¹² In our work for the NIH PROMIS consortium we found that this model applies across all conditions marked by GI symptoms – not just IBS. The current GI symptom framework posits that GI symptoms are captured by 8 domains: (1) Belly Pain; (2) Bloat/Gas; (3) Diarrhea; (4) Constipation; (5) Bowel Incontinence/Soilage; (6) Heartburn/Reflux; (7) Swallowing; and (8) Nausea/Vomiting.

In the absence of finalized PROMIS measures and fully validated symptom indices in IBS, the FDA employs "interim endpoints" that drug manufacturers can use while groups develop new PROs for IBS clinical trials. The interim endpoints apply to both IBS with constipation (IBS-C) and IBS with diarrhea (IBS-D), and measure two aspects of IBS: (1) abdominal pain and (2) abnormal defecation (comprising stool frequency and stool form). Abdominal pain is measured with an 11-point numeric rating scale (NRS), ranging from 0 (no pain) to 10 (worst possible pain). We previously validated the NRS in IBS.¹³ Stool form is measured using the Bristol Stool Scale (BSS), a 7-point index that presents stools spanning the diarrhea to constipation spectrum (Fig. 2). Finally,

DOI of refers to article:

<http://dx.doi.org/10.1016/j.rgmx.2012.10.003>

[☆] See paper by López-Alvarenga JC et al. in page 21-7.

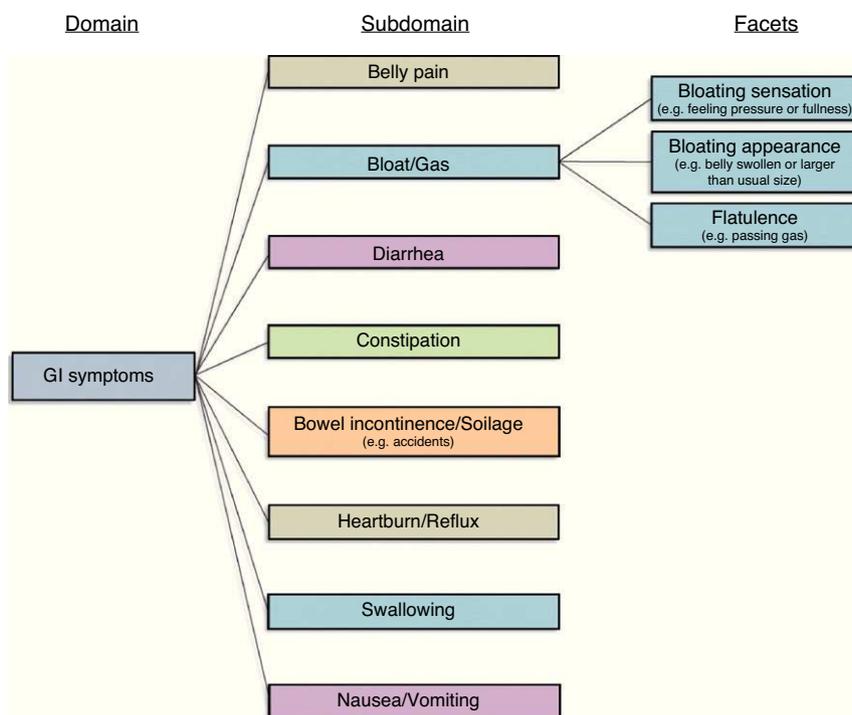


Figure 1 Conceptual model of GI symptoms underlying GI PROMIS item bank development.

stool frequency is measured by patient self-report using a daily stool diary.

According to the FDA, a treatment response in IBS requires simultaneous improvement in abdominal pain and abnormal defecation, both comprising “co-primary” endpoints.¹⁴ In both IBS-C and IBS-D, a pain response is defined by an improvement of $\geq 30\%$ on the NRS when comparing weekly NRS averages over time vs. baseline. Interpretation of a defecatory response varies depending on

the IBS sub-type. For IBS-D, improvement is defined by maintaining an average stool form of ≤ 5 on the BSS for at least 50% of the time. For IBS-C, the FDA requires ≥ 1 complete spontaneous bowel movement (CSBM) per week, regardless of BSS results.

Let’s face it: this is confusing, if not cognitively paralyzing. With the exception of investigators working closely in IBS clinical trials, most clinicians are unaware of these definitions or find them difficult to memorize; they are not particularly intuitive, either. But they serve their role while groups create better endpoints. We expect improved PROs to arrive in the next 1–2 years as groups like the Critical Pathway (C-Path) Institute¹⁵ develop multi-dimensional endpoints for IBS clinical trials.

In this issue of *Revista de Gastroenterología de México*, López-Alvarenga and colleagues present a novel approach to PRO measurement in IBS. Called the “polar vector” method, the approach involves measuring stool frequency and form using a BSS “matrix,” and then converting the matrix data into vectors that track changes in stool frequency and form over time. The approach borrows from basic geometry, which allows calculation of a direction and magnitude of a vector between two points. The mathematics of this approach is outlined in the paper and accompanying technical appendix; I do not review the geometry further in this editorial (quite honestly, I had to open my high school geometry text book to intuit the mathematics of this paper).

To test their concept, the authors employed data from a large, open-label, prospective study using pinaverium bromide and simethicone over four weeks in patients with IBS meeting Rome criteria. The study measured stool consistency and frequency along with improvements in cardinal IBS symptom intensity. Using data from 1,677 patients in the study, the authors created a so-called “omnibus

Bristol stool chart

Type 1		Separate hard lumps, like nuts (hard to pass)
Type 2		Sausage-shaped but lumpy
Type 3		Like a sausage but with cracks on its surface
Type 4		Like a sausage or snake, smooth and soft
Type 5		Soft blobs with clear-cut edges (passed easily)
Type 6		Fluffy pieces with ragged edges, a mushy stool
Type 7		Watery, no solid pieces, Entirely liquid

Figure 2 Bristol Stool Scale.

variable” that accounted for stool consistency (BSS type) and frequency, and employed a “two-dimensional configuration using polar vectors” to display the results. The authors explain that the higher the value of the BSS type and the frequency of the bowel movements, the greater the value of the “polar vector.”

Using this approach, the authors report “polar vector analysis made it possible to show that there was considerable improvement in the IBS-C patients within the first two weeks of treatment and that it remained steady during the final two weeks.” The authors do not define what “considerable improvement” means, or track the changes in polar vector magnitude against patient-reported measures of improvement, such as the Guyatt overall treatment effect (OTE) scale, as recommended by the FDA. Nonetheless, they visually demonstrate changes in the “omnibus variable” using visually dramatic coordinate scales – an unusual, if not spectacular, results graphic for an IBS clinical trial. The investigators further report that a vector magnitude of 12.5 “is apparently equal to a type 4 on the Bristol Stool Scale.” I am unclear precisely how that apparent equation was achieved, but am pleased to see an attempt to lend clinical interpretation to these mathematical concepts. Finally, the authors show how the behavior of the polar vector varied by IBS sub-groups and as a function of the study duration.

In their discussion section, the authors draw wide conclusions regarding the potential benefits of the polar vector approach in IBS PRO measurement. In particular, the authors describe their approach is a “useful method” for evaluating IBS pharmacological therapies, and point out that the approach meets many of the FDA PRO requirements, including a focus on stool consistency and frequency, employing daily symptom diaries, and employing a “multidimensional context.”

Although the approach is quite novel and graphically stimulating, it remains unclear to what degree this technique meets FDA requirements or moves us forward in IBS PRO measurement. First, although the technique may indeed be “useful” to distinguish among IBS subgroups in an uncontrolled, open-label study, the role of this technique in Phase III registration studies remains unclear (as the authors imply in their limitations section). The term “useful” is generally substituted with “valid” in the psychometric literature. For a PRO to be “valid” for the FDA, it must demonstrate face validity (i.e. looks “good” on its face), content validity (i.e. patients support its content through focus groups and cognitive interviews), construct validity (i.e. its scores track with scores of already validated legacy instruments), and criterion validity (i.e. its scores vary meaningfully against the gold standard metric – in this case something like a patient OTE). The polar vector approach does not yet achieve these levels of validity.

In addition, it is unclear whether the approach achieves a “multidimensional” status, as suggested by the authors. A benefit of the approach is its simultaneous capture of both stool frequency and form into one metric. But when the FDA describes a multidimensional PRO, it typically refers to a multi-domain PRO within a broader conceptual framework. In IBS, a multi-domain PRO should indeed not only measure stool frequency and form, but also stool urgency (for IBS-D), straining (for IBS-C), bloating (including how bloating

“looks” vs. “feels”), and pain. The vector approach described in this paper falls short of this multi-domain vision, but is a helpful approach for getting us started.

The success of the polar vector approach depends entirely on the validity of its underlying components – in this case, the BSS itself. The mathematical manipulations of the BSS data in this study are impressive and noteworthy, but they cannot overcome inherent limitations in the underlying data. Although the BSS correlates with intestinal transit time and is widely endorsed as a measure of stool consistency in IBS, there has been surprisingly little work to evaluate IBS patient understanding of the BSS. We previously performed qualitative cognitive de-briefing interviews to solicit patient views about the BSS.¹⁶ For example, we showed the BSS to patients and asked: “*Do you understand what this question is asking you to do?*”

What we found was quite illuminating. In a study with 43 Rome positive IBS patients, we found that most expressed areas of confusion regarding the BSS. 83% noted that a single bowel movement may be characterized by multiple BSS forms (e.g. “Sometimes it will start out very hard and then wind up liquid”), and noted that it would be inaccurate to assign a single consistency to their bowel movement. Many patients noted that they have multiple bowel movements within a single bathroom visit, and that different bowel movements often have different forms. These patients also emphasized that it can be difficult to determine the “start” and “end” of a bowel movement (e.g. “If I get up from the toilet, but then come back a few minutes later, does that ‘count’ as a new BM? Do I assign one form for each BM, or different forms for different bathroom visits?”). 37% noted that their stool consistency varies throughout the day. These patients could not identify a single form to best characterize the day’s bowel movements (e.g. “Am I supposed to give you the average over all my BMs for the day? A typical day involves several BMs with different types;” “My stool in the morning is different than in the evening.”). Patients emphasized that the unit of measurement was unclear (e.g. individual stool vs. bowel movement vs. bathroom visit), and further recommended allowing for separate forms for each bowel movement, and to assign multiple forms *within* bathroom visit.

In short, we found that although the BSS is widely used and endorsed by the FDA and Rome criteria, many patients voice practical concerns about how to respond to this scale. If the BSS is to be included in a future IBS PRO, be it with “polar vectors” or not, its instructions for use will need to be clarified to address pervasive confusion. In addition, the scale itself may require some retrofitting to address its psychometric shortcomings.

Where does this leave us? The polar vector approach is a novel heuristic for visualizing symptom changes in IBS. For that reason alone, it is a useful technique for modifying how we conceive of PRO data, not only in IBS, but also in other gastrointestinal conditions. On the other hand, the polar vector technique is potentially “paralyzing,” insofar as it is difficult to explain, difficult to calculate, and (currently) difficult to interpret clinically using legacy benchmarks. As difficult as the current FDA interim endpoints are to master, the polar vector technique appears to add another layer of complexity atop an already imperfect and oftentimes confusing metric.

In order to make new headway in creating PROs for registration trials, we should break away from paralyzing complexity, and focus principally on PROs built from patient-derived focus groups coupled with FDA best practices for crafting multi-dimensional instruments grounded in a priori conceptual frameworks, as supported by the FDA PRO Guidance document. In the meantime, the polar vector approach gives us a lot to think about and seems worth further investigation to help push us forward in IBS PRO measurement.

Disclaimer

The opinions and assertions contained herein are the sole views of the authors and are not to be construed as official or as reflecting the views of the Department of Veteran Affairs.

References

- Burke LB, Kennedy DL, Miskala PH, et al. The use of patient-reported outcome measures in the evaluation of medical products for regulatory approval. *Clin Pharmacol Ther.* 2008;84:281–3.
- Wagner EH, Austin BT, Von Korff M. Organizing care for patients with chronic illness. *Milbank Q.* 1996;74:511–44.
- Wagner EH, Austin BT, Davis C, et al. Improving chronic illness care: translating evidence into action. *Health Aff (Millwood).* 2001;20:64–78.
- Marshall S, Haywood K, Fitzpatrick R. Impact of patient-reported outcome measures on routine practice: a structured review. *J Eval Clin Pract.* 2006;12:559–68.
- Dobscha SK, Gerrity MS, Ward MF. Effectiveness of an intervention to improve primary care provider recognition of depression. *Eff Clin Pract.* 2001;4:163–71.
- Taenzer P, Bultz BD, Carlson LE, et al. Impact of computerized quality of life screening on physician behaviour and patient satisfaction in lung cancer outpatients. *Psychooncology.* 2000;9:203–13.
- Velikova G, Booth L, Smith AB, et al. Measuring quality of life in routine oncology practice improves communication and patient well-being: a randomized controlled trial. *J Clin Oncol.* 2004;22:714–24.
- Detmar SB, Muller MJ, Schornagel JH, et al. Health-related quality-of-life assessments and patient–physician communication: a randomized controlled trial. *JAMA.* 2002;288:3027–34.
- Reeve BB, Hays RD, Bjorner JB, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the Patient-Reported Outcomes Measurement Information System (PROMIS). *Med Care.* 2007;45:S22–31.
- Cella D, Yount S, Rothrock N, et al. The Patient-Reported Outcomes Measurement Information System (PROMIS): progress of an NIH Roadmap cooperative group during its first two years. *Med Care.* 2007;45:S3–11.
- Spiegel B, Khanna D, Bolus R, et al. Understanding gastrointestinal distress: a framework for clinical practice. *Am J Gastroenterol.* 2011.
- Spiegel BM, Bolus R, Agarwal N. Measuring symptoms in the irritable bowel syndrome: development of a framework for clinical trials. *Aliment Pharmacol Ther.* 2010;32:1275–91.
- Spiegel B, Bolus R, Harris LA, et al. Measuring IBS patient reported outcomes with an abdominal pain numeric rating scale: results from the proof cohort. *Aliment Pharmacol Ther.* 2009.
- Guidance for industry: irritable bowel syndrome — clinical evaluation of drugs for treatment; 2012 [accessed at <http://www.fda.gov/downloads/Drugs/Guidances/UCM205269.pdf>].
- Critical Path (C-Path) Institute; 2013 [accessed at <http://c-path.org/index.cfm>].
- Spiegel BM, Bolus R, Agarwal N, et al. Do patients understand the Bristol Stool Scale? Results of cognitive debriefing of IBS patients. In: Abstract. *Digestive Disease Week.* 2011.

B. Spiegel^{a,b,*}

^a *Department of Gastroenterology, VA Greater Los Angeles Healthcare System, Division of Digestive Diseases, David Geffen School of Medicine at UCLA, United States*

^b *Department of Health Services, UCLA School of Public Health, UCLA/VA Center for Outcomes Research and Education, United States*

*Correspondence address: 11301 Wilshire Blvd, Bldg 115, Room 215, Los Angeles, CA 90073, United States. Tel.: +1 310 268 3256; fax: +1 310 268 4510. E-mail address: bspiegel@mednet.ulca.edu