Editorial

The Patient-Reported Outcomes Measurement Information System (PROMIS): Can We Finally Compare Apples to Oranges?

Abstract: Legacy patient-reported outcome mea`sures lack standardization, resulting in difficulty comparing the results of diverse clinical outcome studies: "You can't compare apples to oranges." To address this concern, the National Institutes of Health initiated the Patient-Reported Outcomes Measurement Information System (PROMIS) to assess common dimensions of a wide range of diseases. PROMIS uses computer adaptive testing: A fluid questionnaire chooses subsequent questions based on the responses to previous questions to efficiently characterize outcomes using only 4 to 6 questions. This greatly reduces survey fatigue. Research correlating PROMIS to legacy measures is of value. For now, some questions may require more information than PROMIS can provide, in which case legacy measures could be preferred. In the future, developing and adding a utility score to PROMIS could assess "value" and allow decision analyses and cost-effectiveness analyses for diverse health interventions. In the end, PROMIS may allow us to compare apples to oranges.

As the saying goes, "You can't compare apples to oranges." This decree has long compelled researchers, editors, and readers to ensure that the conclusions of scientific research are based on sound comparison of meaningful data. Vague and imprecise assessments such as "fair," "good," or "excellent" outcomes, frequently determined or measured by practitioners, have been rightly supplanted by patient-reported outcome measures (PROMs), which have become the quintessential means of comparison for both observational and therapeutic studies. Nevertheless, a criticism of current PROMs is a "lack of standardization" that results in "little comparability among them."[1]

For better or for worse, a countless number of existing PROMs, also described as "legacy" measures, have been used to quantitatively assess function and/or disability. Even if these PROMs are condition- or joint-specific, the myriad and often redundant options result in reporting bias, making it difficult to compare the results of diverse studies. As an example, even if 2 studies use validated measures to assess the treatment of shoulder instability, if one reports outcomes using the Western Ontario Shoulder Instability Index (WOSI) and the other, using the Oxford Shoulder Instability Score (OSIS) . . . well, we can't compare apples to oranges.

The diversity of options is further compounded by the need for reporting disparate measures for each joint (e.g., shoulder) and each respective pathology (e.g., instability vs rotator cuff tear), thus bringing to mind the classic orthopaedic adage, "If you have more than one choice, there may not be a good choice."

Moreover, we currently recommend that PROMs be reported in terms of clinically measurable benefit to our patients and thus specifically in terms of the minimal clinically important difference (MCID, or "Do you notice a change?"), patient acceptable symptom state (PASS, or "Are you satisfied?"), and substantial clinical benefit (SCB, or "Are you a great deal improved?").[2] As a result of the expanding list of PROMs and corresponding designators, it is incredibly difficult to qualitatively compare data across studies, as is performed in systematic reviews, and nearly impossible to quantitatively analyze pooled data in meta-analyses. Apples to oranges!

To address these concerns, in 2004, the National Institutes of Health (NIH) initiated the Roadmap for Clinical Research including the Patient-Reported Outcomes Measurement Information System (PROMIS). PROMIS aims to assess common dimensions of PROMs for a wide range of diseases. In this sense, PROMIS is "generic" rather than specific, yet PROMIS is also divided into primary domains: Physical Health, Mental Health, and Social Health. It is further divided into subdomains, including those most relevant to arthroscopic and related research: Physical Function, Upper Extremity Function, Adult, and Pediatric.[3] As such, PROMIS represents a middle ground between joint- or

condition-specific legacy PROMs and truly generic health measures such as the 36-Item Short Form Survey Instrument (SF-36).

What is particularly promising about PROMIS—we tried to resist but had to use the obvious pun once—is a dynamic design using computer adaptive testing in which a fluid questionnaire chooses subsequent questions based on a patient's responses to previous questions. Using computer adaptive testing, PROMIS efficiently characterizes patient outcomes using only 4 to 6 questions and, to the benefit of both patients and practitioners, greatly reduces survey fatigue in comparison with legacy measures.[4]

The NIH has stated that the beneficial features of PROMIS include the following[3]:

> PROMIS measures have greater precision than most conventional measures. Greater precision (less error) enhances power in a less costly way than increasing sample size.
> PROMIS measures have a larger range of measurement than most conventional measures, decreasing floor and ceiling effects as a result.
> PROMIS measures do all this with fewer items than conventional measures, thereby decreasing respondent burden. When used as computer adaptive tests, PROMIS measures usually require 4 to 6 items for precise measurement of health-related constructs.
> PROMIS measures provide a common metric: the T-score (mean = 50, standard deviation = 10). In most cases 50 equals the mean in the U.S. general population. This metric has also been linked to many other conventional measures, and even if other measures are used, it may be possible to report results on the PROMIS metric, a considerable advantage for ensuring comparability across studies.
> PROMIS measures can be administered alongside [other] measures that assess other aspects of health and function.

However, PROMIS is not without limitations. Some have voiced concerns about the challenges posed by the logistics of readily collecting Web-based data in clinical environments lacking the requisite hardware (or the financial resources) to do so. In addition, concerns persist over the lack of specificity resulting from heightened emphasis on broad domains rather than on specific joints and/or pathologies. PROMIS is also limited with regard to assessment of return to activity including work or sport.

Given these facts, we are left to grapple with a dilemma: PROMIS versus legacy measures? In the long term, we are hopeful that the benefits of PROMIS will outweigh the limitations and ultimately include the paradox of allowing us to successfully compare . . . gulp . . . apples to oranges. In the end, what is most important is the patients' perception of their outcome, and what is most specifically under our influence in this regard is the patients' physical function. The PROMIS Physical Function and Upper Extremity Physical Function PROMs could allow us to measure, report, and compare our results whether we treat a hip, knee, or ankle, or a wrist, elbow, or shoulder . . . apples to oranges!

In the short term, we believe that research comparing and correlating PROMIS to legacy measures is of value. We admire such studies that have been published in *Arthroscopy*.[5-12]

In the future, developing and adding a utility score to PROMIS to assess the "value" of a health state could allow simultaneous assessment of health outcome and utility using a single measure. This could allow researchers to efficiently perform decision analyses and ultimately cost-effectiveness analyses for diverse health interventions. We introduce, and follow, such research with significant interest.[13] As articulated by Arthroscopy Association of North America Past President William R. Beach,[14] "rank and file . . . arthroscopic surgeons . . . must get involved in . . . data collection. . . . (I)ndividual surgeons and groups are going to have to compete for the ability to care for patients based on quality metrics. The question always comes down to what data do we collect."

The NIH may be on the right track. PROMIS may be an efficient solution to PROM data collection and mitigate against reporting bias. Of course, other measures may also be recorded, yet if legacy measures can be shown to correlate with PROMIS, then PROMIS may prevail. For now, context matters. In many cases, PROMIS is appropriate and could be preferred. In other cases, there may be questions that require more information than PROMIS can provide, and disease-specific legacy measures may help us to discern exactly what makes an apple an apple or an orange an orange.

In *Arthroscopy*, PROMIS was unheard of until 2017. Now, scholarly studies and eloquent commentary use and discuss PROMIS in our pages each year. Readers are thus forewarned: Finally, it may be possible to compare apples to oranges!

Michael J. Rossi, M.D., M.S.
Assistant Editor-in-Chief
Andrew J. Sheean, M.D.
Associate Editor
Mark P. Cote, P.T., D.P.T., M.S.C.T.R.
Associate Editor, Statistics
Jefferson C. Brand, M.D.
Assistant Editor-in-Chief
James H. Lubowitz, M.D.
Editor-in-Chief

References

1. National Institutes of Health. Patient-Reported Outcomes Measurement Information System (PROMIS), https://

commonfund.nih.gov/promis/index. Accessed March 1, 2020.

2. Lubowitz JH, Brand JC, Rossi MJ. Our measure of medical research should be appreciable benefit to the patient. *Arthroscopy* 2019;35:1943-1944.

3. Wikipedia Entry, https://en.wikipedia.org/wiki/Patient-Reported_Outcomes_Measurement_Information_System. Accessed March 1, 2020.

4. Health Measures. Intro to PROMIS, http://www.healthmeasures.net/explore-measurement-systems/promis/intro-to-promis. Accessed March 1, 2020.

5. Nwachukwu BU, Beletsky A, Naveen N, et al. PROMIS instruments correlate better with legacy measures in knee cartilage patients at post-operative than at pre-operative assessment [published online February 5, 2020]. *Arthroscopy*. doi: 10.1016/j.arthro.2020.01.036.

6. Shamrock AG, Wolf BR, Ortiz SF, et al. Preoperative validation of the patient-reported outcomes measurement information system in patients with articular cartilage defects of the knee. *Arthroscopy* 2020;36:516-520.

7. Beletsky A, Nwachukwu BU, Manderle BJ, et al. The impact of workers' compensation on patient-reported outcomes measurement information system upper extremity and legacy outcome measures in patients undergoing arthroscopic rotator cuff repair. *Arthroscopy* 2019;35:2817-2824.

8. Makhni EC. Editorial commentary: A primer on the patient-reported outcomes measurement information system—Understanding how this powerful tool can be used for clinical and research gains. *Arthroscopy* 2019;35:775-777.

9. Kahan JB, Kassam HF, Nicholson AD, Saad MA, Kovacevic D. Performance of PROMIS Global-10 to legacy instruments in patients with lateral epicondylitis. *Arthroscopy* 2019;35:770-774.

10. Fabricant PD. Editorial commentary: We PROMIS . . . One more patient-reported outcome questionnaire! *Arthroscopy* 2018;34:1118-1120.

11. Minoughan CE, Schumaier AP, Fritch JL, Grawe BM. Correlation of patient-reported outcome measurement information system physical function upper extremity computer adaptive testing, with the American Shoulder and Elbow Surgeons shoulder assessment form and simple shoulder test in patients with shoulder pain. *Arthroscopy* 2018;34:1430-1436.

12. Fidai MS, Saltzman BN, Meta F, et al. Patient-reported outcomes measurement information system and legacy patient-reported outcome measures in the field of orthopaedics: A systematic review. *Arthroscopy* 2018;34:605-614.

13. Hanmer J, Feeny D, Fischhoff B, et al. The PROMIS of QALYs. *Health Qual Life Outcomes* 2015;13:122.

14. Beach WR. Editorial commentary: Patient reported outcomes measurement information system (PROMIS) may be our promise for the future. *Arthroscopy* 2017;33:1775-1776.