Ann Surg Oncol (2023) 30:3879–3881
https://doi.org/10.1245/s10434-023-13394-7

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

EDITORIAL – RECONSTRUCTIVE ONCOLOGY

Editorial: Maintaining the Integrity of PROMs in Research and Practice

Minji Kim, BS[1], Joline Fong[1], Andrea L. Pusic, MD, MHS[2], John P. Fischer, MD, MPH[3], Babak J. Mehrara, MD[1], and Jonas A. Nelson, MD, MPH[1]

[1]Plastic and Reconstructive Surgery Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, NY; [2]Division of Plastic and Reconstructive Surgery, Department of Surgery, Brigham Health, Boston, MA; [3]Division of Plastic Surgery, Department of Surgery, University of Pennsylvania, Philadelphia, PA

Patient reported outcome measures (PROMs) are questionnaires used in clinical and research settings to better understand patients' perspectives on their conditions and the effect of therapies.[1] The use of PROMs has increased exponentially in the past 10 years due to a broader shift towards a patient-based healthcare system.[2] As such, PROMs are becoming increasingly essential in all facets of medicine, including clinical practice, research, and healthcare funding.[2]

The Food & Drug Administration and Consensus-based Standards for the selection of health status Measurement Instruments (COSMIN) have standard guidelines on developing PROMs and measurement properties to ensure the creation of robust and quality PRO instruments that capture the burden of disease or treatment in its entirety.[3,4] International Society for Quality of Life Research (ISOQOL) also has published recommendations for properly utilizing PROM.[5] These guidelines must be followed accordingly to apply PROMs in clinical and research settings.[3–5] At the minimum, steps needed for PROM development include item generation, item reduction, field testing, and psychometric evaluation.[6] Psychometric properties include reliability and validity, which are necessary to support the use of PROMs.[4] Additional factors may be considered to ensure the PROMs measure a given entity accurately and fairly, including clarity, relevance, and cultural appropriateness and translation.[6] The process also takes into account patients' understanding and interpretation of the items.[7] Widely used PROMs have followed the standard guidelines on developing and validating their measures. As examples, BREAST-Q and M.D. Anderson Symptom Inventory (MDASI) specific for lung cancer (MDASI-LC) underwent a rigorous development methodology and psychometric validation.[7,8]

Many validated PROMs have published guidelines on how to utilize and score their instruments. The European organization for research and treatment of cancer (EORTC)'s Quality of Life and BREAST-Q user manuals are all accessible to researchers and physicians.[9,10] These manuals explicitly indicate that modifications are not allowed as straying from the guidelines can undermine the PROM's reliability, validity, and overall effects.[11] In this editorial, we seek to emphasize the importance of adhering to the specific developer guidelines of PROMs and to provide precautionary measures. As PROMs become more regularly utilized, our goal is to ensure optimal use and accurate measurement of PROMs in both clinical and research settings.

There is a lack of studies on the inappropriate use of PROMs in the current literature, and additional studies investigating inaccurate PROM-related methodologies and results are necessary. However, we can extrapolate from a recent study examining use of BREAST-Q in postmastectomy breast reconstruction. In this study, the most common mistakes in BREAST-Q administration were addition, deletion, or alteration of items, failure to score properly, and completion of the questionnaire through a telephone interview or a third party.[12] Each of these is addressed in the user manual of the BREAST-Q.[10]

Foremost, all questions in a validated PROM should be preserved and be consistent. Adaptation of a questionnaire involves making changes that consider the nuances of the

First Received: 31 January 2023
Accepted: 13 March 2023
Published online: 10 April 2023

J. A. Nelson, MD, MPH
e-mail: Nelsonj1@mskcc.org

language and culture of the patient population. However, modification of a PROM means to change instrument content, format, and/or the mode of administration.[13] Modification may threaten the content validity and psychometric qualities of PROMs, and is thus unacceptable (and generally prohibited by copyright restrictions).[9] Appropriate modification requires development of the new interpretation guideline and comprehensive evidence to support the alterations.[14] Questions should not be added as this can alter how patients respond to other questions.[11] Questions should also not be deleted as missing PRO data can reduce the study power and cause bias interpretation.[15] Finally, questions should not be altered, either in phrasing or wording, as this can minimize the ease of reading, translation, and content coverage.[14] Changing the number and/or phrasing of questions can affects the validity, measurement properties, and the interpretability and integrity of a PROM, as previously noted in an inaccurate use of the Asthma Control Questionnaire.[8, 11] Unnecessary modification also means that new findings cannot be compared to other published data.

There is no one standardized way of scoring PROMs, as each PROM has a unique way of scaling and scoring methods. The variety of scoring methods and interpretation of these scores can complicate accurate scoring.[16] However, an important quality of a PROM is its interpretability, the ability to assign qualitative meaning to a quantitative score or change in the score.[16] Normative or reference values have been defined for many PROMs to put a specific score into perspective and to establish an accurate reference frame.[17] Reference values allow physicians to recognize low scores and provide appropriate and timely intervention, and if PROMs are not scored properly, these reference values can become meaningless. Further, systematic reviews and meta-analyses rely on each individual study to have properly executed its PROM and scoring. For instance, a systematic review on PROMs in lymphedema reviewed 200 related studies which utilized 54 different PROMs, the most common PROMs being SF-36 and EORTC. This review had to trust that each study reviewed adhered to the guidelines and scoring.[18] Systematic reviews or meta-analyses can only synthesize the findings of multiple, independent studies if accurate scoring rules are followed.

Lastly, the essence of a PROM is that patients complete the surveys themselves. The mode of PROM administration is of important consideration and is often specified in the PROM manual. For example, BREAST-Q has been validated to be conducted either in paper or online at the discretion of patient and surgeon. Both means were found to be highly reliable and equivalent.[19] However, some studies administered BREAST-Q via telephone interviews instead, which can introduce a new type of bias.[12] Patients may feel influenced by the interviewer's presence and tone of voice.[20] Whenever possible, PROMs should be self-completed unless manuals' specifications state otherwise or have been validated to be accurate with other methods of administration.

It is a widely understood and often part of the copyright and licensing agreement that providers adhere to PROM guidelines. Assessments should be used as originally intended, without unwarranted modification, to preserve their integrity and validity. Authorized versions of PROMs should be used in clinical and research settings according to developers' intended purpose.[11] As PROMs become more common in clinical and research settings, it is imperative that one preserves the content of PROMs, accurately calculates the score, and administers the PROM as specified. We strongly encourage researchers and physicians to follow the user manuals of all PROMs and refrain from any inappropriate changes to maintain clinical interpretability and to enable improved patient care and subsequent clinical research.

FUNDING N/A

DISCLOSURES John P. Fischer: Consulting for 3M, Baxter, AbbVie, BD, WL Gore, and Integra. Babak Mehrara: Investigator initiated grants from Pfizer, Regeneron, Puretech, Advisor to Mediflix, Royalty payments Puretec.

REFERENCES

1. Voineskos SH, Nelson JA, Klassen AF, Pusic AL. Measuring patient-reported outcomes: key metrics in reconstructive surgery. *Annu Rev Med*. 2018;69:467–79.
2. Churruca K, Pomare C, Ellis LA, et al. Patient-reported outcome measures (PROMs): a review of generic and condition-specific measures and a discussion of trends and issues. *Health Expect*. 2021;24(4):1015–24.
3. Health USDo, Human Services FDACfDE, Research, et al. Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health Qual Life Outcomes*. 2006;4:79.
4. Mokkink LB, Terwee CB, Patrick DL, et al. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol*. 2010;63(7):737–45.
5. Reeve BB, Wyrwich KW, Wu AW, et al. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res*. 2013;22(8):1889–905.
6. Centers for Medicare & Medicaid Services. *Patient-Reported Outcome Measures*. 2022.
7. Mendoza TR, Wang XS, Lu C, et al. Measuring the symptom burden of lung cancer: the validity and utility of the lung cancer module of the M. D. Anderson symptom inventory. *Oncologist*. 2011;16(2):217–27.
8. Cohen WA, Mundy LR, Ballard TN, et al. The BREAST-Q in surgical research: a review of the literature 2009–2015. *J Plast Reconstr Aesthet Surg*. 2016;69(2):149–62.
9. Quality of Life. EORTC Quality of Life: FAQ. https://qol.eortc.org/faq-category/questionnaire-modifications/. Accessed August 15, 2022.

10. *BREAST-Q Version 2.0: A Guide for Researchers and Clinicians.* Memorial Sloan Kettering; 2017.
11. Juniper EF. Validated questionnaires should not be modified. *Eur Respir J*. 2009;34(5):1015–7.
12. Gallo L, Chu JJ, Shamsunder MG, et al. Best practices for BREAST-Q research: a systematic review of study methodology. *Plast Reconstr Surg*. 2022.
13. U.S. Food & Drug Administration. *Principles for selecting, developing, modifying, and adapting patient-reported outcome instruments for use in medical device evaluation.* Rockville, MD 2022.
14. Rothman M, Burke L, Erickson P, Leidy NK, Patrick DL, Petrie CD. Use of existing patient-reported outcome (PRO) instruments and their modification: the ISPOR good research practices for evaluating and documenting content validity for the use of existing instruments and their modification PRO task force report. *Value Health*. 2009;12(8):1075–83.
15. Mercieca-Bebber R, King MT, Calvert MJ, Stockler MR, Friedlander M. The importance of patient-reported outcomes in clinical trials and strategies for future optimization. *Patient Relat Outcome Meas*. 2018;9:353–67.
16. van der Willik EM, Terwee CB, Bos WJW, et al. Patient-reported outcome measures (PROMs): making sense of individual PROM scores and changes in PROM scores over time. *Nephrology (Carlton)*. 2021;26(5):391–9.
17. Hernandez BO, Russell GB, Wright SC Jr, Madden LL. Normative value for the laryngopharyngeal measure of perceived sensation. *Laryngoscope*. 2022;132(2):398–400.
18. Beelen LM, van Dishoeck AM, Tsangaris E, et al. Patient-reported outcome measures in lymphedema: a systematic review and COSMIN analysis. *Ann Surg Oncol*. 2021;28(3):1656–68.
19. Fuzesi S, Cano SJ, Klassen AF, Atisha D, Pusic AL. Validation of the electronic version of the BREAST-Q in the army of women study. *Breast*. 2017;33:44–9.
20. Sinnott CJ, Glickman LT, Natoli NB, Dobryansky M, Veeramachaneni HB. Assessing patient-reported outcomes after female cosmetic genital surgery. *Ann Plast Surg*. 2020;85(S1):S17–22.