Published in final edited form as:

Med Care. 2019 May ; 57(Suppl 5 1): S46–S51. doi:10.1097/MLR.0000000000001086.

Interpreting and Acting on PRO Results in Clinical Practice: Lessons Learned from the PatientViewpoint System and Beyond

Amanda L. Blackford, ScM¹, Albert W. Wu, MD, MPH^{2,3}, and Claire Snyder, PhD^{1,2,3}

¹Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins, 550 N. Broadway, Suite 1111, Baltimore, MD 21205, ablackf1@jhmi.edu

²Johns Hopkins Bloomberg School of Public Health, 624 N. Broadway, Suite, Baltimore, MD 21205, awu@jhu.edu

³Johns Hopkins School of Medicine, 624 N. Broadway, Room 649, Baltimore, MD 21205, Tel: 443-287-5469, csnyder@jhu.edu

Abstract

Introduction: We have used several methods to aid in the interpretation of patient-reported outcome (PRO) scores, and to suggest recommendations for acting on them, both within the PatientViewpoint system and in work done since.

Methods for Interpreting PRO Scores: For identifying possibly concerning change scores, PatientViewpoint uses questionnaires' minimally important differences (MIDs) or score worsening $>1/2$ standard deviation. For poor scores in absolute terms, PatientViewpoint primarily uses distributions based on normative data (e.g., worst quartile, >2 standard deviations from the mean). To advance methods for score interpretation, we explored using needs assessments to identify health-related quality-of-life (HRQOL) scores associated with unmet needs and requiring follow-up. We also investigated the ability of PRO scores to predict patients' most bothersome issues.

Methods for Addressing PRO Results: To develop suggestions for addressing issues identified by PRO questionnaires, we conducted a targeted literature review, interviewed experts from different disciplines, developed draft recommendations based on the literature and interviews, and finalized the recommendations in a consensus meeting with all experts.

Discussion: The needs-assessment method requires strongly correlated content in the needs assessment and HRQOL questionnaire. Additional research is needed to explore using the worst scores in absolute terms to identify patients' most bothersome issues. The approach described for developing suggestions for addressing PRO results focuses on local resources and is best-suited for a local context.

Corresponding Author: Amanda L. Blackford, ScM.

Disclosures:

A. Blackford has received funding from Genentech for the work on this paper and has no other conflicts to disclose.

A. Wu has received personal consulting fees from Glaxo Smith Kline, ViiV, and Gilead, and has no other conflicts to disclose.

C. Snyder has received funding from Genentech for the work on this paper and royalties from UpToDate and has no other conflicts to disclose.

INTRODUCTION

Starting in 2005, a multi-disciplinary team developed an electronic system for collecting patient-reported outcome (PRO) data and linking the results with the electronic health record for clinical use.¹ The resulting system, PatientViewpoint (www.PatientViewpoint.org), was designed as a “questionnaire delivery device” to serve as a repository of questionnaires, from which clinicians can “order” questionnaires much in the same way as they order laboratory tests or imaging studies. Clinicians can decide which patients will complete which questionnaires, and at what intervals.

PatientViewpoint underwent two phases of pilot-testing in breast and prostate cancer patients at the Johns Hopkins Cancer Center. Phase 1 established the feasibility and value of using PatientViewpoint in clinical practice.² Phase 2 used a continuous quality improvement approach to refine the system’s design and operations.³ PatientViewpoint has continued to be used in the breast cancer program to collect PROs for a registry of patients on hormonal therapy⁴ and as part of a randomized trial of a remote weight-loss intervention.⁵

In developing PatientViewpoint, a number of issues were raised. These included determining which scores should be highlighted for clinicians’ attention, as well as whether and how to advise clinicians on addressing issues identified by the PRO questionnaires. In terms of interpreting scores, PRO results in PatientViewpoint are reported using line graphs of scores over time for each of the measured domains, and possibly concerning scores are highlighted in yellow. However, we needed to determine which score ranges would be considered “possibly concerning.”

In terms of acting on the PRO results, PatientViewpoint includes a “What Can I Do” link on the clinicians’ score report for each domain. Clinicians can click the link to access recommendations for acting on the PRO results. However, we needed to develop the recommendations that would be provided.

To address the questions raised in PatientViewpoint’s development regarding interpreting and acting on the PRO scores, we explored a number of methodologic approaches. In this paper, we report on the methods we have used to address the issues of PRO score interpretation and acting on PRO results – both in PatientViewpoint, and in subsequent research.

METHODS FOR AIDING SCORE INTERPRETATION

A Pragmatic Approach...and Beyond

When initially developing PatientViewpoint, we used a pragmatic approach to identify possibly concerning scores requiring a clinician’s attention. Where possible, we used established minimally important differences (MID) for the various questionnaires. For example, for the European Organization for Research and Treatment of Cancer (EORTC) Quality-of-Life Questionnaire-Core 30 (QLQ-C30), worsening of 10 points or greater was highlighted based on published data.^{6–8} For questionnaires in PatientViewpoint without established MIDs, we used 1/2 a standard deviation change, based on evidence that that is a

fair estimate of the MID.⁹ To identify scores that were poor in absolute terms, we primarily relied on comparisons to normative data and highlighted scores that were, for example, in the worst quartile or 2 standard deviations from the mean. We referred to the published literature and/or questionnaire reference manuals for normative data and standard deviation estimates.

While these pragmatic approaches were workable, we sought more evidence-based data regarding which scores should be highlighted for clinical attention. Specifically, score distributions do not really indicate at which point a PRO score becomes possibly concerning – or whether the score represents an issue that is bothersome for the patient. Given that in clinical practice, we are most interested in highlighting possibly concerning and/or bothersome issues, we conducted a series of analyses exploring how to identify which scores are considered concerning or bothersome.

We used two approaches: 1) using needs assessments to identify scores on health-related quality of life (HRQOL) questionnaires associated with an unmet need, and 2) exploring whether HRQOL domain scores can predict issues bothering patients the most. The focus on HRQOL, and in particular the EORTC-QLQ-C30,¹⁰ emerged from previous research that supported its use for monitoring patients in clinical practice compared to a needs assessment.¹¹

The EORTC-QLQ-C30 includes five functional domains (physical, role, emotional, social, cognitive), eight symptoms (fatigue, pain, nausea/vomiting, appetite loss, constipation, diarrhea, insomnia, dyspnea), global health/QOL, and financial impact. With the exception of the global health/QOL questions, which use a 7-point rating scale, responses are on a 4-point scale from ‘not at all’ to ‘very much.’ Items within a domain are summed and transformed to a continuous 0 to 100 scale. Higher scores represent better function or worse symptoms. Thus, unlike a needs assessment for which response categories separate patients into those with and without an unmet need, the QLQ-C30’s scores do not have an inherent meaning – which is to say, it is unclear at what point on the 0–100 scale a score becomes possibly concerning or bothersome. We therefore sought to identify scores requiring a clinician’s attention.

Needs Assessments to Identify Cut-offs

Objective—The objective of our first set of analyses explored whether needs assessments could be used to identify HRQOL scores representing an unmet need, and therefore, requiring a clinician’s attention. Using data from the U.S., Japan, Canada, and Denmark, we explored whether two different needs assessments (the Supportive Care Needs Survey-Short Form-34 [SCNS-SF34] and the Three-Levels-of-Needs-Questionnaire [3LNQ]) could identify possibly concerning scores on the QLQ-C30.^{12–15} Another study, conducted in the Netherlands, also looked at relationships between the EORTC head-and-neck specific module and the SCNS-Head and Neck Cancer.¹⁶

The SCNS-SF34 covers five domains of health needs: physical and daily living, psychological, sexual, patient care and support, and health system and information.^{17–18} Patients respond using a one-month recall period on a 5-point scale: 1=‘not applicable’,

2='satisfied', 3='low unmet need', 4='moderate unmet need', 5='high unmet need'. The 3LNQ was developed to measure needs using three different approaches: problem intensity, problem burden, and felt need.¹⁹ Patients answer questions covering a range of issues from physical activity, sexual function, and pain to worrying, social problems, and depression. Responses range from 'not at all' to 'very much'.

Methods—First, we dichotomized the needs assessments scores into two categories: unmet need vs. no unmet need. For the SCNS-SF34, items in a domain are averaged, so domain scores > 2.0 represent an unmet need. On the 3LNQ, the 'need for pain relief' item was used to group patients into unmet need vs. no problem/met need/no wish.

Second, we qualitatively associated the QLQ-C30 domains with needs assessment items or domains based on the similarity of the content. The goal was to link the continuously measured HRQOL domains with the needs assessment item/domain likely to have the strongest association.

To test the level of the association, we calculated the area under the receiver operating characteristic (ROC) curve (AUC) for each hypothesized item/domain. This method is commonly used to describe the performance of the potential classifier.²⁰ An AUC of 0.50 represents chance classification, and an AUC of 1.0 represents perfect classification. We only pursued further analysis for QLQ-C30 domains associated with SCNS-SF34 items/domains with AUC .70. If there were multiple items/domains with AUC .70, we selected the item with the greatest AUC, indicating the strongest relationship.

It may be useful to perform simple descriptive analyses on the matched HRQOL domains and needs assessment items/domains with AUC .70. For example, we examined the distribution of the QLQ-C30 scores, separately for patients with and without the unmet need, through boxplots. Formal tests for differences in the means or medians could be conducted with t-tests (for normally distributed scores) or Wilcoxon rank sum tests (for skewed distributions), respectively. Scores could also be categorized using quartiles or tertiles, and statistical tests such as Fisher's exact or Chi-square tests, depending on the number of observations in each cell of the cross-tabulation, can describe the overall association.

Third, we examined the ROC curve and selected one or more potential cutoffs for maximizing sensitivity (probability of detecting true positives) and specificity (probability of detecting true negatives).

Results—The results of these analyses have been published previously.^{12–16} Briefly, our hypotheses regarding which QLQ-C30 domains were likely to have a strong association with an SCNS-SF34 domain/item were largely confirmed, with for example, the QLQ-C30 domain of physical function having a strong association with the SCNS-SF34 item "work around the home" and emotional function with "feelings of sadness". However, QLQ-C30 domains such as appetite loss and constipation did not have a specific match on the SCNS-SF34. In the 3LNQ study, the pain domain from the QLQ-C30 was associated with the 'need for pain relief.'

When examining the ROC curve to select a cut-off, we considered the trade-offs between sensitivity and specificity. For example, the ROC curve in Figure 1 has an AUC of 0.81, and the cutoffs of the QLQ-C30 scores that yield each possible sensitivity and specificity are shown along the curve. Table 1 presents two potential cutoffs: 80 and 90. Using a cutoff of 80 yields sensitivity of 65% (17/26) and specificity of 83% (70/84). The positive predictive value (PPV) and negative predictive value (NPV) of this cutoff are equal to 55% (17/31) and 89% (70/79), respectively. Using a cutoff of 90 results in a more sensitive (85%), but less specific (58%), test with PPV=39% and NPV=92%. Notably, our analyses conducted using data from the U.S., Japan, Canada, and the Netherlands produced generally consistent results, both in terms of the SCNS-SF34 item with the strongest association to the QLQ-C30 domain, and in terms of the sensitivity and specificity of the various cutoffs.^{12–14,16}

Predicting Patients' Most Bothersome Issues

Objective—Another approach we explored was to determine whether HRQOL scores predicted which issues were bothering patients the most, and if so whether a patient's worst score at a given time point vs. the greatest worsening in score from a previous time point predicted these bothersome issues more accurately.²¹ This analysis represents an alternative approach of determining which PRO scores should be highlighted for a clinician's attention based on patient-perceived bother. This study used scores from the QLQ-C30 and patients' separate reports of up to two most bothersome issues related to function and to symptoms. Patients completed the questionnaires at up to three different time points.

Methods—First, we examined the data for missingness, number of issues identified, and extreme observations. We removed patients with missing data at any time point, observations where there were no reported or more than two reported bothersome issues in either function or symptoms, and observations where a patient had perfect scores (100 for function, and 0 for symptoms) across all domains.

We then selected the two QLQ-C30 function domains furthest from 100 and the two symptom domains furthest from 0 (i.e., the 'worst' scores). In cases of ties, we selected all domains with the worst score. For patients who completed the questionnaires at consecutive visits, we calculated change scores (follow-up minus previous time point's score). The two domains demonstrating the largest positive change for symptoms and largest negative change for functioning, indicating deterioration, were considered the 'worst' change scores. Patients demonstrating no worsening at the follow-up time point were excluded. We selected absolute and changes scores for analysis because it is often unclear whether the patient's current issues (absolute score) or relative change from the last visit is more informative of their current need.

Next, we created an indicator variable for each of the two function and symptom issues the patient flagged as most bothersome. For each most-bothersome domain, we assessed whether it was also selected as having the worst score for that patient and calculated the percentage of patients with agreement (accuracy). This same approach was then applied to the change scores to determine agreement. The accuracy between worst scores vs. change scores were qualitatively compared. Note that change scores tend to be more variable than

absolute scores, so any analyses comparing these measures directly before categorizing them should take this into account.

Other variations in the analytic approach include simultaneously examining agreement for patients who indicated two issues as most bothersome. The percentage of patients for whom the worst score correctly identified both issues raised, only one issue, or neither issue can be calculated. The same can be done for the change score, and these measures of accuracy can be qualitatively compared to each other.

Results—The results have been published previously.²¹ Briefly, the poorest scores in absolute terms better predicted the issues bothering patients the most. For example, among patients reporting one bothersome issue, absolute scores identified it correctly 91% of the time for both function and symptoms (vs. 40% and 26%, respectively, for change scores).

METHODS FOR DEVELOPING RECOMMENDATIONS FOR ACTING ON PRO RESULTS

Objective

In addition to our efforts to aid PRO score interpretation in PatientViewpoint, we conducted a multi-stage, multi-disciplinary process to develop suggestions for how to act on possibly concerning scores for each of the PRO domains.²² The goal was for the recommendations to be brief enough to appear in a pop-up box.

Methods

First, we determined the PRO domains covered in the questionnaires of interest. Because PatientViewpoint houses multiple PRO questionnaires and because there may be overlap in the domains covered in different PRO questionnaires, we aimed to make the recommendations generic across questionnaires (i.e., not developing separate recommendations for the physical function domain of the QLQ-C30 and for the SCNS-SF34).

Second, we conducted a targeted literature review to identify interventions to address issues in each of the domains of interest. In our case, we conducted a targeted MEDLINE search, referred to key palliative care textbooks and websites, and searched reference sections for additional sources. We summarized the literature-based findings regarding potential actions to take in response to issues in each of the domains.

Third, we conducted in-person, one-on-one interviews with experts from multiple disciplines to learn how they would respond to problems with each of the PRO domains in their own context. We consulted with experts in palliative care, medical oncology, radiation oncology, internal medicine, oncology social work, psychiatric liaison nursing, triage nursing, clergy, patient advocacy, and outcomes research. The research team oriented the interviewee to the goals of the project and the PRO domains of interest, with the specific questionnaires available for review. Interviewees were asked to think about patients with issues in these PRO domains and to discuss possible causes and means to address the issues, including both

what they could do, as well as the role of other members of the care team. The specific context and resources available for our institution were considered in the expert's recommendations. The recommendations from each interview for each domain were summarized.

Fourth, we held an in-person consensus development meeting bringing all the experts together. At the beginning of the meeting, we reviewed the project's goals and demonstrated how the final recommendations would be integrated in PatientViewpoint. Based on the literature- and expert-based recommendations, the research team had developed draft recommendations. For each domain, the draft recommendation was projected on the screen, and the experts had the opportunity to add, delete, or modify content (with the changes being made in real-time) until consensus was reached. The goal was not to provide one specific action to take in response to an issue, but to offer a range of options for consideration. Experts could refer to a document that summarized the recommendations for each PRO domain from the literature and from each interview. We finalized consensus recommendations for 20 PRO domains in 2 hours.

Results

The results of this process have been published in detail elsewhere.²² Briefly, the recommendations generally included options related to assessment, referral, lifestyle modification, or pharmacologic therapy. These final recommendations were built into PatientViewpoint and can be accessed in the clinicians' score report by clicking on a "What Can I Do" link for each domain. Thus, clinicians logged into PatientViewpoint can access these recommendations in real-time when reviewing a patient's score report; however, this information is not yet available in our electronic health record (EHR). As part of our continuous quality improvement evaluation, we asked clinicians for feedback on this process. While feedback was generally positive, clinicians suggested that more detailed referrals to specific resources (e.g., the specific phone numbers to call) would be helpful for What Can I Do, and also strongly endorsed the integration of the PRO results within the EHR.³

DISCUSSION

When implementing PROs in clinical practice, it is important to help clinicians interpret what the scores mean, and to provide actionable suggestions for clinicians to consider. In this paper, we reviewed both our initial pragmatic approaches, as well as follow-on analyses we conducted to help identify possibly concerning scores. The methods presented use statistical approaches that are relatively simple to implement with straightforward steps from preliminary data setup and exploration to formal model estimation and summarization of results.

For the needs assessment approaches, ROC analysis can be applied to any continuously measured outcome to assess a binary classification (need vs. no need). Using sensitivity and specificity to assess the performance of a potential predictor is a well-known approach that translates well into the clinical setting. However, the cutoffs determination relied on the ability to match specific HRQOL domains with the needs assessment's content, and some

domains did not have good matches. In addition, when using PRO questionnaires in this way, one needs to be cognizant of common method bias²³. Further, if the patient sample is too homogenous, there may be insufficient variability in the scores to evaluate a range of potential cutoffs, leading to large deviations in sensitivity and specificity from cutoff to cutoff. Clinically, it is important to consider the relative importance of false positives or false negatives. While the previous studies found general consistency across the four different national groups in terms of both matching needs items with HRQOL domains and the cutoff scores identified, it is possible that some cutoffs may not be generalizable to different patient populations.

We conducted only one analysis comparing the worst absolute score vs. the greatest worsening. While the initial findings support the worst scores in absolute terms as the preferred approach for predicting issues bothering patients the most, additional studies are needed to confirm this finding. These studies will require both longitudinal HRQOL assessment, as well as patients' reports of their most bothersome issues. Data from patients with no bothersome issues cannot be used in analyses, so applying this method to a relatively healthy or high-functioning population may not be useful.

Finally, strengths of our approach for developing suggestions for acting on PRO results include the multi-stage, multi-disciplinary methods, which included input both from a targeted literature review as well as from experts from a variety of backgrounds. We aimed to develop recommendations that were domain-specific rather than questionnaire specific, such that the suggestions would apply regardless of which questionnaire was used to assess a particular outcome. However, our approach specifically focused on the local resources available in our institution, so the recommendations are not necessarily generalizable to other settings. Other institutions aiming to develop recommendations should consider the trade-offs between developing suggestions for their specific context vs. more generic and generalizable recommendations. Our recommendations were developed for clinicians specifically, and alternative approaches could focus on patient self-management advice.

As the use of PROs in clinical practice becomes increasingly common, methods to aid in the interpretation of scores and help clinicians act on the results are critical. When we first deployed PatientViewpoint, we had to rely on pragmatic methods based on limited evidence. Because of this gap, we have developed and tested a range of approaches that can help inform understanding of PRO scores' meaning and possible approaches for addressing issues identified by the PRO questionnaires. To summarize our findings, using needs assessments to identify possibly concerning scores on HRQOL questionnaires has worked effectively in multiple studies. This method can be used in cases where there is a close linkage between the content covered in the HRQOL questionnaire and needs assessment. In addition, one of our analyses suggests that the poorest HRQOL scores in absolute terms are more likely to identify the issues bothering patients the most compared to the greatest worsening; however, this finding requires replication before it can be implemented. Our approach to developing What Can I Do works best in a local context. A modified approach would be needed to develop more generalizable recommendations.

Acknowledgments

The research summarized in this paper was supported by the National Cancer Institute (1R21CA134805-01A1; 1R21CA113223-01) and American Cancer Society (MRSG-08-011-01-CPPB). Dr. Snyder is a member of the Sidney Kimmel Comprehensive Cancer Center at Johns Hopkins (P30CA006973). This paper is part of the PRO-cision Medicine Methods Toolkit paper series funded by Genentech. The PRO-cision Medicine Methods Toolkit paper series was presented during a symposium at the 2018 Annual Conference of the International Society for Quality of Life Research (Dublin, Ireland).

REFERENCES

1. Snyder CF, Jensen R, Courtin SO, et al. PatientViewpoint: A website for patient-reported outcomes assessment. *Qual Life Res* 2009;18:793–800. [PubMed: 19544089]
2. Snyder CF, Blackford AL, Wolff AC, et al. Feasibility and value of PatientViewpoint: A web system for patient-reported outcomes assessment in clinical practice. *Psychooncology* 2013;22:895–901. [PubMed: 22544513]
3. Wu AW, White SM, Blackford AL, et al. Improving an electronic system for measuring PROs in routine oncology practice. *J Cancer Surviv* 2016;10:573–582. [PubMed: 26644190]
4. Smith KL, Yeruva SLH, Blackford A, et al. Predictors of adherence to adjuvant endocrine therapy for early breast cancer in a prospective clinic-based cohort. *San Antonio Breast Cancer Symposium* December 5–9, 2017 San Antonio, Texas.
5. Santa-Maria CA, Coughlin J, Blackford A, et al. POWER-remote: A randomized study evaluating the effect of a remote-based weight loss program on biomarkers in women with early stage breast cancer. *San Antonio Breast Cancer Symposium* December 6–10, 2016 San Antonio, Texas.
6. Osoba D, Rodrigues G, Myles J, et al. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998;16:139–144. [PubMed: 9440735]
7. King MT. The interpretation of scores from the EORTC quality of life questionnaire QLQ-C30. *Qual Life Res* 1996;5: 555–567. [PubMed: 8993101]
8. Maringwa JT, Quinten C, King M, et al. Minimal important differences for interpreting health-related quality of life scores from the EORTC QLQ-C30 in lung cancer patients participating in randomized controlled trials. *Support Care Cancer* 2011;19:1753–1760. [PubMed: 20886240]
9. 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;41:582–592. [PubMed: 12719681]
10. Aaronson NK, Ahmedzai S, Bergman B, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993;85:365–376. [PubMed: 8433390]
11. Snyder CF, Herman JM, White SM, et al. When Using Patient-Reported Outcomes in Clinical Practice, the Measure Matters: A Randomized Controlled Trial. *J Oncol Pract*. 2014;10:e299–306. [PubMed: 24986113]
12. Snyder CF, Blackford AL, Brahmer JR, et al. Needs assessments can identify scores on HRQOL questionnaires that represent problems for patients: An illustration with the Supportive Care Needs Survey and the QLQ-C30. *Qual Life Res* 2010;19:837–845. [PubMed: 20340048]
13. Snyder CF, Blackford AL, Okuyama T, et al. Using the EORTC QLQ-C30 in clinical practice for patient management: Identifying scores requiring a clinician's attention. *Qual Life Res* 2013;22:2685–2691. [PubMed: 23532341]
14. Snyder CF, Blackford AL, Sussman J, et al. Identifying changes in scores on the EORTC-QLQ-C30 representing a change in patients' supportive care needs. *Qual Life Res* 2015;24:1207–1216. [PubMed: 25398495]
15. Johnsen AT, Petersen MA, Snyder CF, et al. How does pain experience relate to the need for pain relief? A secondary exploratory analysis in a large sample of cancer patients. *Support Care Cancer* 2016;24:4187–95 [PubMed: 27169702]
16. Jansen F, Snyder CF, Leemans R, Verdonck-de Leeuw IM. Identifying cut-off scores for the EORTC QLQ-C30 and the Head and Neck Cancer Specific Module (EORTC QLQ-H&N35)

representing unmet supportive care needs in head and neck cancer patients. *Head Neck*. 2016;38 Suppl 1:E1493–500 [PubMed: 27111696]

17. Bonevski B, Sanson-Fisher RW, Girgis A, et al. Evaluation of an instrument to assess the needs of patients with cancer. *Cancer* 2000;88:217–225. [PubMed: 10618626]
18. Sanson-Fisher R, Girgis A, Boyes A, et al. The unmet supportive care needs of patients with cancer. *Cancer* 2000;88:226–237. [PubMed: 10618627]
19. Johnsen AT, Petersen MA, Pedersen L, et al. Development and initial validation of the Three-Levels-of-Needs Questionnaire for self-assessment of palliative needs in patients with cancer. *J Pain Symptom Manage* 2011;41:1025–1039. [PubMed: 21306865]
20. Hosmer DW, Lemeshow S. *Applied logistic regression* (2nd edition). Chichester, NY: Wiley; 2000.
21. Snyder CF, Blackford AL, Aaronson NK, et al. Can patient-reported outcome measures identify cancer patients' most bothersome issues? *J Clin Oncol* 2011;29:1216–1220. [PubMed: 21343558]
22. Hughes EF, Wu AW, Carducci MA, et al. What Can I Do? Recommendations for responding to issues identified by patient-reported outcomes assessments used in clinical practice. *J Support Oncol* 2012;10:143–148. [PubMed: 22609239]
23. Podsakoff PM, MacKenzie SB, Lee JY, and Podsakoff NP. Common method biases in behavioral research: A critical review of the literature and recommended remedies. *J Applied Psych* 2003; 88:879–903.

KEY TAKE-HOME POINTS

- Using needs assessments to identify possibly concerning scores on HRQOL questionnaires has worked effectively in a number of studies. It is a viable approach for aiding interpretation in cases where there is a needs assessment with content strongly correlated with the HRQOL questionnaire domains of interest.
- While initial analyses suggest that patients' worst scores in absolute terms, compared to the greatest worsening from the previous time point, are better predictors of the issues bothering them the most, further research is needed to confirm this finding.
- Using a multi-stage, multi-disciplinary process to develop recommendations for addressing issues with PRO domains can produce a range of options to inform management. Clinics need to consider the trade-offs between tailoring the recommendations to locally relevant resources, or making more generic recommendations that would be more generalizable. We addressed only recommendations for clinicians, and patient self-management advice could also be considered.

A combination of pragmatic solutions and exploratory research can inform interpreting and acting on PRO scores.

Figure 1.
ROC curve for the discriminative ability of the physical function domain of the QLQ-C30 in predicting patients who have trouble performing work around the home.

Table 1.

Example of a cross-tabulation of SCNS needs by classification of needs according to QLQ-C30 scores using two different cutoffs.

		QLQ-C30				
		Cutoff = 80		Cutoff = 90		Total
		Classified Negative	Classified Positive	Classified Negative	Classified Positive	
SCNS	True Negative True Positive	70 9	14 17	49 4	35 22	84 26
	Total	79	31	53	57	110