Patient-Reported Outcomes to Inform Tolerability Characterization in Cancer Trials: A Critical Perspective on Unresolved Methodological Issues

Bill Byrom, PhD[1] and John Devin Peipert, PhD[2,3]

Abstract

Patient-reported outcomes assessing treatment toxicity and its impact, and aspects of quality of life, have become increasingly important components of oncology drug development including early phase dose optimization studies. Despite insights from these data, there are limitations on drawing inferences about whether a treatment is tolerable or intolerable, which requires understanding individual risk-benefit assessments that vary between patients, cancer types, and stages of disease. This commentary examines unresolved methodological issues requiring more research rigor and attention: the development of tolerability thresholds based upon toxicity severity data; the direct measurement of tolerability through specific patient self-report measures; and overcoming the challenge of baseline value assessment in important measures such as single item side-effect bother assessments. We call for further research to resolve these challenges and drive more robust inference-making from patient-reported outcomes data collected in cancer clinical trials.

Plain Language Summary

Improving How We Measure Treatment Tolerability in Cancer Trials: When new cancer treatments are being tested, researchers need to understand not just whether they work (e.g., do they remove tumors or reduce their size), but also whether patients can tolerate them. Current clinical trials use questionnaires to ask patients about the severity of treatment side effects and how much they impact daily activities. This is valuable to understand treatment toxicity and to enable comparison of the side effect profiles of different drugs. But, this doesn't tell us whether patients are able to tolerate these side effects given the potential benefits of treatment.

This commentary argues that we need better methodologies to translate the reported severity of side effects to an understanding of tolerability, by developing threshold values for severity scores to define tolerable and intolerable treatment, based on patient insights. We also argue that it is important to explore direct measures of tolerability, perhaps by asking patients if they would be willing to continue treatment. Additionally, some measurement issues need to be resolved, such as asking about side effects before treatment has begun.

Addressing these methodological gaps will ensure that we can derive accurate findings based on the important patient-reported outcomes data collected throughout cancer drug development.

[1]eCOA Science, Signant Health, Blue Bell, PA, USA
[2]Department of Applied Health Sciences, University of Birmingham, Birmingham, UK
[3]Centre for Patient Reported Outcomes Research, University of Birmingham, Birmingham, UK

Corresponding author:
Bill Byrom, PhD, eCOA Science, Signant Health, 785 Arbor Way, Blue Bell, PA 19422, USA.
Email: bill.byrom@signanthealth.com

Patient-reported outcomes have become an increasingly important component of oncology drug development programs, measuring, amongst other things, toxicity, adverse events, the impact of side effects, quality of life, mental health, and cognition. Much of this has been driven by renewed emphasis on patient-focused drug development, with the US Food and Drug Administration (FDA) catalyzing thinking through its four-part guidance series.[1] In cancer trials, the importance of implementing appropriate patient-reported outcome measures and following good measurement timing and strategy is underlined in disease-specific industry guidance,[2] along with the drive to include the direct patient perspectives in helping to characterize tolerability in optimal dose finding in early phase development.[3,4]

Despite this, methodological issues remain that must be the focus of ongoing research to ensure that even greater value can be derived from the data that patients faithfully report during clinical trials of new treatments. Continuing to optimize our methods will help ensure patient-reported outcomes add value for patients, clinicians, regulators, and health technology assessment (HTA) bodies. In this short commentary, we touch on unresolved methodological areas related to the use of patient-reported outcomes to help characterize the tolerability of oncological treatments, identifying existing and emerging approaches and future work that is important to consider to improve the value of the data and the robustness of inference and decision making.

Tolerability: Are We Measuring It?

Tolerability is defined by the International Council for Harmonisation as "the degree to which overt adverse effects can be tolerated by the subject".[5] Recognizing some limitations in this definition, the Friends of Cancer Research convened an expert team that broadened this definition to more explicitly identify the patient's role in reporting on tolerability, noting that, "[…] A complete understanding of tolerability should include direct measurement from the patient on how they are feeling and functioning while on treatment".[6] The Patient Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE) is a system of measures that enable patients to report their experience of (usually) a subset of the total 78 symptomatic adverse events (AEs).[7] Most AEs are scored by severity ("none" to "very severe"), with some AEs containing additional items measuring frequency ("never" to "almost constantly") or interference ("not at all" to "very much"). While interference begins to explore the impact of an AE on daily activities, it is only available for 23/78 (29.5%) of the AEs included. PRO-CTCAE data has been reported in the label for BLENREP® (belantamab mafodotin, GSK) in characterizing increased severity of blurred vision during treatment.[8]

We argue that PRO-CTCAE provides a description of toxicity, and not specifically whether the toxicity experienced was tolerable or intolerable. Tolerability involves self-appraising risk-benefit by the individual patient made in the context of their own life's circumstances. Cancer patients may be willing to tolerate far higher levels of toxicity compared to patients with other diseases. Moreover, two individuals with the same cancer diagnosis may tolerate the same treatment differently. One qualitative study illustrated this risk-benefit dynamic through one patient's narrative: "If it continues to recede … obviously I'm going to put up with all the symptoms and side effects I get in favor of having to get rid of the cancer".[9] Single summary measures derived from PRO-CTCAE data, such as the toxicity index,[10] can be used to provide a measure of combined burden of multiple toxicities. This composite measure, however, provides only one approach on how these data should be combined to calculate a single metric, may fail to weight multiple toxicities correctly, lacks clear clinical meaning and interpretability, and while quantifying toxicity still fails to distinguish "tolerable" from "intolerable".

The FACT-G GP5 single item measuring the bother of side effects ("not at all" to "very much")[11] is a valuable measure to assess the combined impact of AEs. The measure has been shown to be related to early treatment discontinuation, often a sign of lack of treatment tolerability.[12] This has been further demonstrated in the FDA approved label for RETEVMO® (selpercatinib, Eli Lilly) which states that "patient reported overall side effect impact results were supported by a lower incidence of treatment discontinuation due to adverse reactions".[13] Qualitative research with cancer patients has indicated that the GP5 item can flexibly represent the impact of multiple side effects or a single most burdensome side effect troubling an individual patient, including the overall impact on their ability to participate in routine activities such as work.[14] This flexibility suggests the GP5 item's responsiveness to individual patients' life circumstances that may underpin risk/benefit considerations is related to tolerability, though it does not capture this consideration directly.

Quality of life measures can provide insight into the impact of treatment side effects, such as the physical function and role function subscales of the EORTC QLQ-C30.[15] Still, the challenge remains on how these scores, based on impact scales ("not at all" to "very much"), can determine the willingness of a patient to tolerate continued treatment. They measure an aspect of tolerability, but not tolerability itself.

So, measuring toxicity and its impact, the measures above provide insights into aspects of tolerability and enable direct comparison of the side effects and impacts of different treatments, providing valuable information to guide treatment decision making. However, they fail to directly answer the question of whether a treatment is tolerable or intolerable in the context of a specific disease. To answer this question, we need to consider the meaningfulness of the toxicity profile and its impact, an area needing more research, as discussed below.

Understanding Tolerability From Patient-Reported Toxicity and Treatment Impact Data

PRO-CTCAE data are commonly analyzed using a combination of descriptive and statistical modeling techniques. Descriptive statistics provide summaries of symptom frequency and severity, with visualization methods like line and bar charts illustrating trends and distributions over time.[16] More formal statistical methods, such as repeated-measures ANOVA and generalized estimating equations, model the change in symptom ratings over time, adjusting for within-patient variation and covariates such as treatment arm and baseline status.[16] Time-to-event analyses, such as Kaplan-Meier estimates, are valuable in measuring the time to first appearance or worsening of symptoms above a predefined threshold (e.g., severity rating ≥ 3).[16] Area under the curve calculations can be used to describe longitudinal toxicity profiles of individual AEs,[16] and composite scores (including the toxicity index[10] mentioned above) aim to quantify the combined toxicity of multiple AEs into a single measure.[17] More recently, exposure–response analysis has been identified as another valuable approach, specifically in early-phase trials where pharmacokinetic data are available.[18] These methods can evaluate the relationship between drug exposure levels and patient-reported AEs, providing insights into how varying drug concentrations may affect the frequency, severity, or onset of toxicities captured by PRO-CTCAE, supporting dose optimization decision making.[19]

These approaches become more valuable if we understand the threshold scores or ranges that define when side effects become intolerable for different cancer populations, accounting for individual risk-benefit considerations. One approach to defining tolerability thresholds for single measures, such as when evaluating a single AE or a composite of multiple AEs, is to anchor against harder observable endpoints such as treatment pauses, dose reductions, or early discontinuation due to side effects. Meta analyses of data from existing trials may provide a good starting point to quantify this threshold value. Alternatively, qualitative methods such as evaluating the tolerability of hypothetical combinations of AEs (vignettes)[20] may help to define tolerability thresholds with multidimensional data such as the PRO-CTCAE. These qualitative and anchor-based approaches are analogous to well understood techniques to determine meaningful change thresholds for other clinical outcome assessment measures. Tolerability thresholds will enable us to both characterize toxicity and infer its associated tolerability. More research is encouraged on methods to translate toxicity measures into tolerability threshold values.

Direct Measures of Tolerability

Development of direct measures of tolerability should be explored. Simply asking "would you be willing to continue this treatment as is?" alongside the GP5 and PRO-CTCAE might be a starting point to enable a patient to weigh observed or anticipated efficacy against toxicity. This question would provide a direct and patient-centered assessment of treatment tolerability, integrating patients' values, expectations, and experiences into a single actionable metric, rather than relying on translating toxicity reports into tolerability inferences. Willingness questions like this may help bridge the gap between the clinical significance of toxicities and their practical impact on daily life, supporting more nuanced dose optimization and clinical decision-making, especially in early-phase trials.

However, willingness to continue treatment is very likely to also be influenced by dispositional factors (patient characteristics that exist before treatment begins) beyond toxicity, such as disease severity, treatment expectations and hopes, perceived efficacy, and individual resilience.[21] These factors contribute to substantial inter-patient variability, making it challenging to use toxicity data alone to predict individual tolerability, underlining the importance of including direct assessments of tolerability in addition to characterizing toxicity and its impact in clinical drug trials. However, we do not recommend measuring these dispositional variables in themselves within a measure of tolerability. While explaining variability between individuals, these additional items do not deliver meaning that can be associated with attributes of treatment and aid regulatory decision making. We similarly observe this same phenomenon in the estimation of meaningful change thresholds for other clinical outcome assessments in which dispositional factors add variance to individual notions

of meaningful improvement and worsening. Tolerability is multifaceted and formal qualitative work in patients to develop a direct measure to be used alongside toxicity assessment is encouraged, and this may lead more towards a multi-item measure for defining and measuring tolerability in cancer patients.

A Final Conundrum: Baseline Assessment Using GP5

While collection of the GP5 at baseline is generally well accepted, especially as most endpoints consider a change from baseline score, how do patients interpret a question about side effects of treatment before they actually receive the study treatment?

Patients entering a trial subsequent to previous treatment may report some bother from side effects associated with the previous treatment.[22] This can make it difficult to understand the true degree of bother associated with the new treatment when considering a change from baseline endpoint. Further, it has been suggested that treatment naïve patients may report non-zero baseline bother values based on negative expectations of the treatment they are to receive, or fears about potential known side effects.[14] One analysis of three solid tumor clinical trials reported 11.8–15.7% of cancer treatment-naïve patients, and 23.9% of treatment-experienced patients reported bother scores of 2-4 ("somewhat" to "very much") at baseline.[23]

Further, some studies have reported a high proportion of missing GP5 assessments at baseline which may reflect difficulties in interpreting the item pre-treatment. However, measuring baseline health status relative to the symptoms a patient is already experiencing is important to reduce potential bias in evaluating and characterizing treatment-related tolerability. This allows clinicians and researchers to distinguish treatment-emergent adverse effects from stable pre-existing symptoms, and supports better interpretation of change over time. To limit misunderstandings related to item interpretation and skipping of GP5 at baseline, it is recommended that additional instructional text is developed and validated with patient input to accompany the measure when administered at the baseline or at any other pre-treatment assessment. For example, instructional text such as "If you are currently receiving treatment for your cancer, please rate the side effects of current treatment you have experienced over the past 7 days, rather than the side effects you expect or anticipate having. If you are not currently receiving treatment, for each question please rate your current symptoms over the past 7 days, and if absent select 'None'" could be developed and tested for understanding and meaning in patients.

Conclusion

Patient-reported outcomes have become indispensable in oncology clinical trials, yet critical methodological gaps remain that limit the value we can derive from the data patients provide. This commentary has highlighted fundamental challenges that demand urgent attention from the research community.

First, we must understand how to best leverage toxicity reports to make inferences about tolerability. This requires concerted effort to establish meaningful thresholds that distinguish tolerable from intolerable treatment experiences across different cancer populations. We call on the research community to establish studies that anchor PRO-CTCAE, GP5, and quality of life measures against observable tolerability indicators, and to explore qualitative methodologies to help define multidimensional tolerability thresholds across different cancer patient populations.

Second, the development and validation of direct tolerability measures should be accelerated. Simple willingness-to-continue questions may offer immediate utility, but formal qualitative research is needed to develop robust, multi-item measures that capture the true complexity of tolerability from the patient perspective.

Third, the baseline assessment challenge for measures such as the GP5 requires more evaluation. We urge patient involvement in the development of standardized instructional text that clarifies item interpretation pre-treatment, ensuring baseline data can reliably inform change-from-baseline analyses.

Through collaborative, systematic methodology research we can ensure that patient-reported outcome measures collected in cancer clinical trials fulfill their promise of truly patient-centered drug development.

ORCID iDs

Bill Byrom https://orcid.org/0000-0002-9186-4271
John Devin Peipert https://orcid.org/0000-0001-5762-7881

Author Contributions

The draft and final versions of this article were conceptualized, designed, written and edited by both authors. BB and JDP conceived and wrote this article.

Funding

Declaration of Conflicting Interests

Data Availability Statement

Data sharing not applicable to this article as no datasets were generated or analyzed.

References

1. US Food and Drug Administration. *FDA Patient-Focused Drug Development Guidance Series for Enhancing the Incorporation of the Patient's Voice in Medical Product Development and Regulatory Decision Making*. FDA. https://www.fda.gov/drugs/development-approval-process-drugs/fda-patient-focused-drug-development-guidance-series-enhancing-incorporation-patients-voice-medical. Updated March 21, 2025. Accessed October 17, 2025.

2. US Food and Drug Administration. *Core Patient-Reported Outcomes in Cancer Clinical Trials: Guidance for Industry*. FDA; 2024. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/core-patient-reported-outcomes-cancer-clinical-trials. Accessed October 17, 2025.

3. US Food and Drug Administration. *Project Optimus*. FDA. https://www.fda.gov/about-fda/oncology-center-excellence/project-optimus. Updated June 12, 2024. Accessed October 17, 2025.

4. US Food and Drug Administration. *Optimizing the Dosage of Human Prescription Drugs and Biological Products for the Treatment of Oncologic Diseases: Guidance for Industry*. FDA. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/optimizing-dosage-human-prescription-drugs-and-biological-products-treatment-oncologic-diseases. Updated August 8, 2024. Accessed October 17, 2025.

5. European Medicines Agency. *ICH E9 statistical principles for clinical trials - Scientific guideline*. EMA; 1998. https://www.ema.europa.eu/en/ich-e9-statistical-principles-clinical-trials-scientific-guideline. Accessed October 17, 2025.

6. Basch E, Campbell A, Hudgens S, et al. *Broadening the Definition of Tolerability in Cancer Clinical Trials to Better Measure the Patient Experience*; 2018. https://friendsofcancerresearch.org/wp-content/uploads/Comparative-Tolerability-Whitepaper_FINAL.pdf. Accessed: October 16, 2025.

7. Basch E, Reeve BB, Mitchell SA, et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). *J Natl Cancer Inst*. 2014;106(9):dju244. doi:10.1093/jnci/dju244.

8. US Food and Drug Administration. *BLENREP® (belantamab mafodotin-blmf) for injection, for intravenous use*. FDA. https://www.accessdata.fda.gov/drugsatfda_docs/label/2025/761440s000lbl.pdf. Updated October, 2025. Accessed January 05, 2026.

9. Martin ML, Correll J, Walding A, Rydén A. How patients being treated for non-small cell lung cancer value treatment benefit despite side effects. *Qual Life Res*. 2022;31(1):135-146. doi:10.1007/s11136-021-02882-6.

10. Gresham G, Diniz MA, Razaee ZS, et al. Evaluating Treatment Tolerability in Cancer Clinical Trials Using the Toxicity Index. *J Natl Cancer Inst*. 2020;112(12):1266-1274. doi:10.1093/jnci/djaa028.

11. FACIT Group. *FACIT Item GP5*. FACIT; 1987. https://www.facit.org/measures/facit-item-gp5. Accessed October 17, 2025.

12. Peipert JD, Zhao F, Lee JW, et al. Patient-Reported Adverse Events and Early Treatment Discontinuation Among Patients With Multiple Myeloma. *JAMA Netw Open*. 2024;7(3):e243854. doi:10.1001/jamanetworkopen.2024.3854.

13. US Food and Drug Administration. *RETEVMO® (selpercatinib) capsules, for oral use*. FDA. https://www.accessdata.fda.gov/drugsatfda_docs/label/2024/218160s002s003lbl.pdf. Updated September, 2024. Accessed October 17, 2025.

14. Peipert JD, Shaunfield S, Kaiser K, et al. How do patients interpret and respond to a single-item global indicator of cancer treatment tolerability? *Support Care Cancer*. 2023;31:37. doi:10.1007/s00520-022-07484-7.

15. European Organisation for Research and Treatment of Cancer (EORTC). *Quality of Life in Cancer Patients*. EORTC; 1995. https://qol.eortc.org/questionnaire/eortc-qlq-c30/. Accessed October 17, 2025.

16. Regnault A, Loubert A, Gorsh B, et al. A toolbox of different approaches to analyze and present PRO-CTCAE data in oncology studies. *J Natl Cancer Inst*. 2023;115(5):586-596. doi:10.1093/jnci/djad018.

17. Mead-Harvey C, Basch E, Rogak LJ, et al. Statistical properties of items and summary scores from the Patient-Reported Outcomes version of the Common Terminology Criteria for Adverse Events (PRO-CTCAE®) in a diverse cancer sample. *Clinical Trials*. 2024;22(2):161-169. doi:10.1177/17407745241286065.

18. Overgaard RV, Ingwersen SH, Tornøe CW. Establishing Good Practices for Exposure-Response Analysis of Clinical Endpoints in Drug Development. *CPT Pharmacometrics Syst Pharmacol*. 2015;4(10):565-575. doi:10.1002/psp4.12015.

19. Xia H, Booth BP, Wang Y, et al. Use of patient-reported outcomes (PRO) data to complement exposure-response analysis in early clinical cancer drug development. *J Patient Rep Outcomes*. 2023;7(1):116. doi:10.1186/s41687-023-00651-2.

20. Staunton H, Willgoss T, Nelsen L, et al. An overview of using qualitative techniques to explore and define estimates of clinically important change on clinical outcome assessments. *J Patient Rep Outcomes*. 2019;3(1):16. doi:10.1186/s41687-019-0100-y.

21. Peipert JD, Smith ML, On Behalf of the EVOLV Study Team. Reconsidering tolerability of cancer treatments: opportunities to focus on the patient. *Supportive Care in Cancer*. 2022;30(5):3661-3663. doi:10.1007/s00520-021-06700-0.

22. Roydhouse JK, King-Kallimanis BL, Roy P, et al. Exploration of baseline patient-reported side effect bother from cancer therapy. *Clinical Trials*. 2020;17(3):332-337. doi:10.1177/1740774520910389.

23. Roydhouse J, Breslin M, Zola A, et al. Patient-Reported Side Effect Bother: Understanding the Value of the Baseline Report. *Patient*. 2026;19(1):93-101. doi:10.1007/s40271-025-00766-2.