Why PROMs are failing regulatory scrutiny, and how to fix them

Written by The Life Science Feed Editorial Team · Edited by Mara Voss

> The shift towards patient-centric drug development has elevated patient-reported outcome measures (PROMs) from supplementary data to endpoints that determine drug approval. Regulators now expect PROMs to demonstrate not just statistical significance, but also robust psychometric properties and clinical meaningfulness. This elevated bar means many PROMs, particularly those developed without rigorous methodology, are simply not fit for purpose.

KEY TAKEAWAYS

- The Pivot: Regulatory bodies now demand PROMs meet stringent psychometric and clinical meaningfulness criteria, moving beyond mere statistical significance.
- The Data: A PROM must demonstrate content validity, reliability, construct validity, and responsiveness to change to be considered acceptable.
- The Action: Developers must engage with regulatory guidance early and invest in rigorous PROM development and validation from the outset of clinical programs.

Patient-reported outcome measures are designed to capture the patient's perspective on their health status, symptoms, and functional limitations. These instruments are invaluable for assessing treatment benefit in areas where objective clinical measures may not fully capture the impact on daily life, such as pain, fatigue, or quality of life. But their subjective nature also makes them susceptible to bias and measurement error, which regulators are increasingly keen to identify and eliminate.

The development of a new drug often hinges on demonstrating a meaningful benefit to patients. When that benefit is primarily subjective, a well-constructed PROM becomes the primary evidence. This means the PROM itself must be as rigorously developed and validated as any other clinical endpoint. Without this foundational work, even a statistically significant change in a PROM score may be dismissed as uninterpretable by regulatory agencies.

The Foundational Requirements for PROMs

Regulators expect PROMs to adhere to a set of core psychometric properties. The first and arguably most critical is content validity. This means the PROM must comprehensively cover all relevant aspects of the patient's experience of the disease and its treatment, as understood by the target patient population. If a PROM misses key symptoms or impacts that patients care about, it fails at this initial hurdle. This often requires extensive qualitative research, including patient interviews and focus groups, to ensure the instrument's items are relevant and understandable.

Reliability is another non-negotiable requirement. A reliable PROM consistently produces the same results under the

same conditions. This includes test-retest reliability, where scores are stable over time in stable patients, and internal consistency, where items within a scale measure the same underlying construct. If a PROM's scores fluctuate wildly without a corresponding change in the patient's condition, its utility as a measure of treatment effect is severely compromised.

Construct validity assesses whether the PROM actually measures what it purports to measure. This is often evaluated by examining how the PROM correlates with other measures that theoretically should be related (convergent validity) or unrelated (discriminant validity). For example, a quality of life PROM should correlate positively with other established quality of life measures and negatively with measures of disease severity. Without strong construct validity, the interpretation of changes in PROM scores becomes speculative.

Finally, a PROM must demonstrate responsiveness to change. This means the instrument must be sensitive enough to detect clinically meaningful improvements or deteriorations in a patient's condition. A PROM that cannot differentiate between a stable patient and one who has genuinely improved or worsened is of little value in a clinical trial. This is often assessed by comparing changes in PROM scores with changes in objective clinical markers or global impression scales.

Defining Clinical Meaningfulness

Beyond the psychometric properties, regulators demand evidence of clinical meaningfulness. A statistically significant change in a PROM score does not automatically translate into a benefit that matters to patients. This is where the concept of a 'minimally important difference' (MID) comes into play. The MID is the smallest change in a PROM score that patients perceive as beneficial or detrimental. Establishing the MID is a complex process, often involving anchor-based methods (correlating PROM changes with external indicators of change) and distribution-based methods (using statistical properties of the PROM scores).

Without a clearly defined and justified MID, regulators may struggle to interpret the clinical relevance of a treatment effect. A drug that achieves a statistically significant but clinically insignificant change in a PROM score is unlikely to gain approval. This is particularly true for chronic conditions where small, incremental improvements may be important to patients, but the threshold for regulatory acceptance remains high. The growing focus on patient experience in areas like weight management highlights the need for PROMs that capture subtle but impactful changes.

The Development Process and Regulatory Engagement

The journey of a PROM from concept to regulatory acceptance is arduous. It typically begins with a thorough literature review and qualitative research to define the conceptual framework and generate items. These items are then refined through cognitive debriefing with patients to ensure clarity and cultural appropriateness. The resulting draft instrument undergoes pilot testing and psychometric validation in a relevant patient population.

Early and continuous engagement with regulatory bodies is paramount. Developers should seek scientific advice on their PROM strategy, including the choice of instrument, its validation plan, and the proposed interpretation of results. This proactive approach can identify potential pitfalls early, saving considerable time and resources. Ignoring regulatory guidance until late-stage development is a common reason for PROM-related setbacks.

But the challenges extend beyond initial validation. PROMs must also be translated and culturally adapted for use in multinational trials, a process that requires careful linguistic and conceptual equivalence. A direct translation is rarely

sufficient; the instrument must resonate with patients in different cultural contexts. This is a process, often requiring multiple rounds of translation, back-translation, and cognitive debriefing.

Where PROMs Fall Short

Many PROMs fail regulatory scrutiny because they are not developed with the same rigor as other clinical endpoints. Some are generic instruments, not specific enough to capture the specificities of a particular disease. Others are developed by clinicians without sufficient patient input, leading to instruments that measure what clinicians think is important, rather than what patients actually experience. This disconnect is a frequent source of regulatory concern.

Another common issue is the lack of a clear conceptual framework. A PROM should be grounded in a theoretical model that explains how the disease impacts patients and how treatment is expected to modify these impacts. Without this framework, the selection of items can be arbitrary, and the interpretation of scores becomes difficult. This is particularly relevant when considering the long-term impact of therapies, a topic often explored in discussions around chronic disease management.

The statistical analysis of PROM data also presents challenges. While traditional statistical methods are often applied, the ordinal nature of many PROM scales requires careful consideration. The choice of statistical methods for analyzing change scores, handling missing data, and defining responder analyses can all influence regulatory acceptance. Regulators are increasingly sophisticated in their review of PROM data, demanding robust statistical justification for all analytical choices.

The open-label design of some trials can also introduce bias into PROM reporting. Patients who know they are receiving an active treatment may report greater improvements than those who are blinded, even if the objective benefit is similar. While blinding is not always feasible, trial designs must account for this potential bias, perhaps through the use of independent, blinded assessors for certain endpoints or by incorporating placebo run-in periods.

The regulatory environment for PROMs continues to evolve. As more drugs target symptoms and quality of life, the demand for high-quality, validated PROMs will only increase. Developers who invest in rigorous PROM methodology from the outset, and who engage proactively with regulatory agencies, will be better positioned to demonstrate the true value of their therapies to patients and secure market authorization. The alternative is a costly and time-consuming cycle of re-analysis and re-submission, often ending in disappointment. For a deeper dive into the complexities of clinical trial design, the Harrison's Principles of Internal Medicine offers comprehensive insights into the evidence base for various interventions.

> **CLINICAL IMPLICATIONS**
>
> Clinicians often rely on a drug's label to understand its full benefit, but if the patient-reported outcomes supporting that label are weak, the true impact on daily life remains ambiguous. This means that even approved drugs may have an unclear real-world benefit for symptoms that matter most to patients, forcing clinicians to rely on anecdotal evidence or their own subjective assessments.
>
> For drug developers, the message is clear: PROMs are not an afterthought. Investing in their rigorous development and validation is as critical as any preclinical or Phase I study. Skimping on this front will lead to costly delays or outright rejection, regardless of a drug's objective efficacy on surrogate markers.
>
> Patients, meanwhile, are increasingly vocal about their desire for treatments that improve their quality of life,

> not just their lab values. When PROMs fail regulatory muster, it is often because they do not adequately capture what patients truly value. This disconnect can lead to therapies that are technically effective but ultimately unsatisfying for those who need them most.

EDITORIAL & AI STANDARDS

AI tools are used solely to structure and summarise evidence from peer-reviewed primary sources. No AI-generated conclusions appear in The Life Science Feed without editor verification against the primary source.

REFERENCES

1. 1. Saha P, Cady-McCrea C, Puvanesarajah V, Mesfin A. Patient-Reported Outcomes for Spine Oncology: A Narrative Review. World Neurosurg. 2024;185:165-170. doi:10.1016/j.wneu.2024.02.042
2. 2. Almeida A, Martins T, Lima L. Patient-Reported Outcomes in Sarcoma: A scoping review. Eur J Oncol Nurs. 2021;50:101897. doi:10.1016/j.ejon.2021.101897
3. 3. Tan A, Glass DA 2nd. Patient-reported outcomes for keloids: a systematic review. G Ital Dermatol Venereol. 2019;154(2):148-165. doi:10.23736/S0392-0488.18.06089-3
4. 4. D'Amico E, Haase R, Ziemssen T. Review: Patient-reported outcomes in multiple sclerosis care. Mult Scler Relat Disord. 2019;33:61-66. doi:10.1016/j.msard.2019.05.019
5. 5. Tan YY, Chin YR, Bajaj K, Thomas S. Patient-reported outcomes for nipple reconstruction: Review of literature. Surgeon. 2021;19(5):e245-e255. doi:10.1016/j.surge.2020.11.005
6. 6. Nagendrababu V, Vinothkumar TS, El-Karim I, et al. DENTAL PATIENT-REPORTED OUTCOMES IN ENDODONTICS - A NARRATIVE REVIEW. J Evid Based Dent Pract. 2023;23(1):101805. doi:10.1016/j.jebdp.2022.101805

LICENCE & RIGHTS

MEDICAL DISCLAIMER

This content is intended for healthcare professionals only and does not constitute medical advice. Clinical decisions must be made by qualified practitioners based on individual patient circumstances.

FOLLOW THE LIFE SCIENCE FEED

Instagram @lifesciencefeed **LinkedIn** linkedin.com/company/thelifesciencefeed

thelifesciencefeed.com