JOURNAL OF ABDOMINAL WALL SURGERY
EUROPEAN HERNIA SOCIETY

TYPE Opinion
PUBLISHED 12 November 2026
DOI 10.3389/jaws.2026.17292

Check for updates

OPEN ACCESS

*CORRESPONDENCE
Jacob Rosenberg,
jacob.rosenberg@regionh.dk

RECEIVED 03 July 2026
REVISED 24 October 2026
ACCEPTED 03 November 2026
PUBLISHED 12 November 2026

CITATION
Rosenberg J, Reistrup H, Gram-Hanssen A and Baker JJ (2026) When most patients do well: the case for the patient acceptable symptom state alongside mean differences in surgical patient-reported outcomes.
J. Abdom. Wall Surg. 5:17292.
doi: 10.3389/jaws.2026.17292

COPYRIGHT
© 2026 Rosenberg, Reistrup, Gram-Hanssen and Baker. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

When most patients do well: the case for the patient acceptable symptom state alongside mean differences in surgical patient-reported outcomes

Jacob Rosenberg[1]*, Hugin Reistrup[1], Anders Gram-Hanssen[1,2] and Jason Joe Baker[1]

[1]Department of Surgery, Center for Perioperative Optimization, Copenhagen University Hospital – Herlev and Gentofte, Herlev, Denmark, [2]Gastro Unit, Surgical Section, Copenhagen University Hospital – Amager and Hvidovre, Hvidovre, Denmark

KEYWORDS
ceiling effect, clinically meaningful outcomes, incisional hernia repair, mean difference, patient acceptable symptom state

Introduction

Patient-reported outcome measures (PROMs) in benign procedures such as hernia repair are logical and important [1]. However, interpreting the clinical relevance of PROM scores remains a major methodological challenge, and there is no universally accepted approach to determining when differences in PROM scores are clinically meaningful [2]. Registry infrastructure now enables us to report PROM scores for cohorts of thousands of patients using digital infrastructure for data collection [3]. While this provides considerable statistical power, the choice of reporting strategy can markedly influence the interpretation of the results. In large samples, even a small between-group difference in mean PROM scores may reach statistical significance despite limited clinical relevance. This problem is particularly evident when most patients report perfect outcomes with the highest possible PROM scores, while a smaller group experiences persistent symptoms of varying severity. In this situation, the mean score is dominated by the large group with excellent outcomes, so clinically meaningful differences among patients with ongoing symptoms translate into only marginal changes in the overall average. The result is a pronounced ceiling effect, where mean differences are difficult to interpret despite statistical significance. In this common situation, we argue that the Patient Acceptable Symptom State (PASS) is a more interpretable and useful outcome than the difference in means, and we illustrate the point with adjusted data from a large linked survey and registry study of incisional hernia repair.

What PASS measures

Rather than focusing on small differences in mean PROM scores, PASS focuses on the number of patients who have reached an outcome they consider acceptable. Originally introduced in rheumatology [4, 5] and now widely used in orthopedic surgery [6], the concept defines the PROM threshold above which patients consider their state satisfactory: the level at which a patient would say, "I feel well enough." It reframes the analysis from "how many points apart are the groups?" to "what proportion of patients are in an acceptable state?" The summary statistic becomes the proportion achieving PASS, and the between-

group contrast becomes a difference in proportions: figures that any clinician, patient, or policymaker can grasp at once. Importantly, the threshold is anchored to patients' own judgment rather than to an arbitrary statistical cut.

The same data, two stories

The difference between the two approaches can be illustrated using adjusted data from patients undergoing incisional hernia repair (unpublished data from the AFTERHERNIA Project [7]). We compared patients with small hernia defects (0–2 cm) with those with larger defects (6–10 cm). We used the Abdominal Hernia-Q (AHQ) [8] as a PROM. Two AHQ-based outcomes were examined: the AHQ sum score (16–64 points) and the AHQ physical impact score (6–24 points). The AHQ physical impact score reflects the physical impact subscale within the overall AHQ sum score. Higher scores indicate better outcomes for both measures. For this illustrative example, patients scoring below the 75th percentile of the respective AHQ score distributions were classified as not achieving PASS, corresponding to an AHQ sum score <54 and an AHQ physical impact score <23. These distribution-based thresholds were used solely for illustrative purposes, as established anchor-based PASS thresholds for the AHQ have not yet been established. Although anchor-based thresholds are preferable, the present example illustrates how reporting the proportion of patients above or below a specified threshold may complement comparisons of mean PROM scores.

Using the AHQ sum score (Figure 1), the adjusted proportion of patients not achieving PASS was 25.9% among those with 0–2 cm defects and 33.4% among those with 6–10 cm defects. The adjusted difference was 7.5 percentage points (95% CI 2.8 to 12.3; P = 0.002), corresponding to approximately 7 to 8 additional patients per 100 with larger defects who did not reach an acceptable symptom state. By comparison, the adjusted mean AHQ sum score was 56.8 points in the 0–2 cm group and 55.2 points in the 6–10 cm group, an adjusted mean difference of −1.5 points (95% CI −2.4 to −0.7; P < 0.001).

A similar pattern was observed for the AHQ physical impact score (Figure 1). The adjusted proportion of patients not achieving PASS was 22.9% in the 0–2 cm group and 30.4% in the 6–10 cm group, corresponding to an adjusted difference of 7.6 percentage points (95% CI 2.9 to 12.2; P = 0.001). However, the adjusted mean physical impact score differed by only −0.4 points (95% CI −0.6 to −0.1; P = 0.010), from 22.9 points in the 0–2 cm group to 22.5 points in the 6–10 cm group.

Thus, the PASS-based estimates told a consistent and clinically intuitive story in both scoring approaches: patients with larger defects were about 7–8 percentage points less likely to reach an acceptable symptom state. In contrast, the mean-score interpretation depended strongly on the scale used: the same clinical comparison appeared as a 1.5-point difference on the 16–64 AHQ sum score scale, but only a 0.4-point difference on the 6–24 physical impact scale.

The arithmetic of a ceiling effect

A ceiling effect makes standard summary statistics difficult to interpret. In the data shown here, AHQ scores clustered near the upper end of their respective scales. On the AHQ sum score, adjusted mean scores were 56.8 and 55.2 on a scale from 16 to 64. On the physical impact score, adjusted mean scores were 22.9 and 22.5 on a scale from 6 to 24. These values indicate generally good outcomes in both defect-width groups, but they do not clearly convey the proportion of patients who still had an unacceptable symptom state.

The problem is not only that mean differences are small. It is also that their apparent size depends on the scoring range. A difference of −1.5 points on the AHQ sum score and −0.4 points on the physical impact score may appear quite different, even though the PASS-based analyses showed nearly identical differences in the proportion of patients below an acceptable symptom state. Although expressing mean differences relative to their respective scale ranges improves comparability, it does not directly convey the proportion of patients reaching an acceptable symptom state, which is the principal advantage of PASS-based outcomes. This illustrates a practical limitation of mean PROM scores: they require the reader to understand the scale, the direction, the range, and the likely clinical meaning of a given point difference. By contrast, the PASS-based result is immediately interpretable. A 7–8 percentage-point difference means that, for every 100 patients with larger defects rather than smaller defects, approximately 7 to 8 more patients did not reach a patient-acceptable symptom state.

When most patients report high scores, clinically important variation often lies in the lower tail of the distribution, among patients who do not achieve a good outcome. Mean scores summarize the entire distribution and are therefore dominated by the majority of patients who do well. PASS instead focuses directly on the clinically relevant question: how many patients have not reached an acceptable state? For this reason, PASS should not necessarily replace mean PROM scores, but it can provide a more clinically interpretable summary when ceiling effects are present.

Discussion

An experienced reader will object that dichotomizing a continuous outcome is poor practice; it discards information and usually reduces statistical power [9]. The objection is sound, but it targets a different practice: the arbitrary splitting of continuous variables (median splits, "high versus low") chosen for convenience and often driven by the data at hand. PASS differs in three ways. First, the threshold is externally derived and patient-anchored, fixed before the analysis rather than chosen to fit the data. Second, under a strong ceiling, the upper range of the continuous score carries little usable information anyway, so the information sacrificed by dichotomizing is small. Third, the resulting outcome, the proportion of patients in an acceptable state, is exactly the interpretable quantity the mean fails to deliver. The honest recommendation is therefore not "PASS instead of the mean" but "PASS alongside the mean." When the two agree on significance yet diverge in meaning, the PASS-based estimate should anchor the interpretation.

PASS is not a universal threshold and should be applied thoughtfully. The PASS threshold is specific to the instrument,

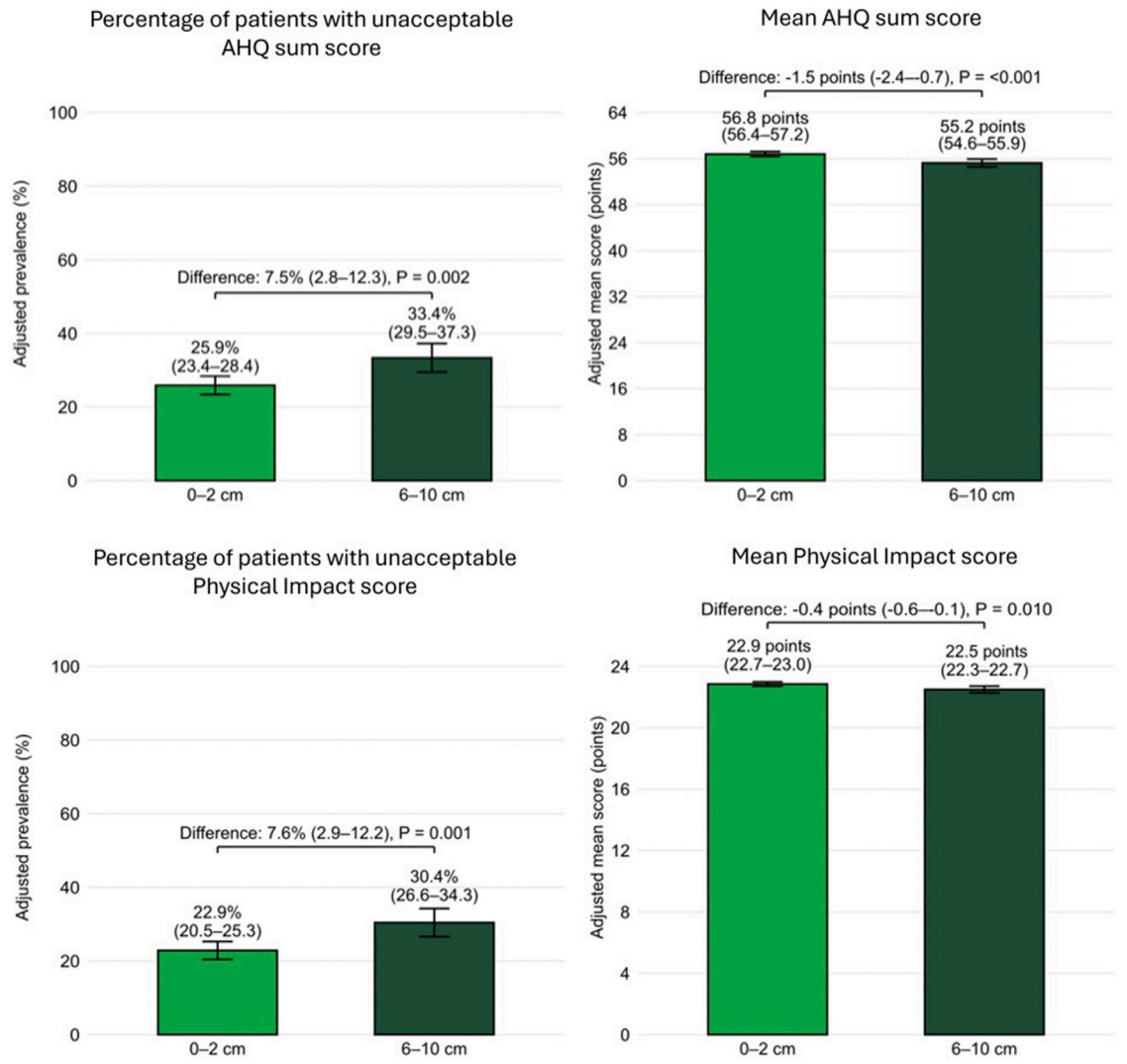


FIGURE 1
Adjusted mean scores and proportions not achieving PASS by hernia defect width. The graphs show Abdominal Hernia-Q (AHQ) sum scores, ranging from 16 to 64, and physical impact scores, ranging from 6 to 24. Bar charts present adjusted estimates for patients with incisional hernia and defect width 0–2 cm or 6–10 cm. Analyses included only responders. For each score, the left panel shows the adjusted proportion of patients not achieving a patient acceptable symptom state (PASS), and the right panel shows the adjusted mean score. Estimates were obtained using g-computation and adjusted for age, sex, Charlson Comorbidity Index, surgical approach, follow-up time, and mesh/closure technique. Error bars indicate 95% confidence intervals. Differences compare patients with defect width 6–10 cm with those with defect width 0–2 cm. PASS not achieved was defined as an AHQ sum score <54 and an AHQ physical impact score <23.

the subscale, and the population; a value derived for one PROM or setting cannot be borrowed for another. Authors should state how their PASS was obtained, typically as an anchor-based estimate or as the 75th percentile of scores. Importantly, the thresholds used in our illustrative example were distribution-based rather than anchor-based and should therefore not be interpreted as validated PASS thresholds. They serve solely to illustrate the potential value of reporting proportions alongside mean PROM scores. Ideally, the threshold is taken from, or validated in, an independent sample rather than estimated from the same data that drive the primary analysis, to avoid optimistic cut-points. Although survey and registry cohorts of this size can absorb the power cost of dichotomization, the PASS analysis should still be pre-specified. Finally, PASS should not be confused with the minimal important difference (MID) [10]: the MID concerns the smallest difference worth noticing, whereas PASS concerns an acceptable absolute state. The two answer different questions and are complementary [6].

Much of modern hernia and benign-surgery outcomes research studies populations that, by design, mostly do well. In these settings, differences in mean PROM scores are easy to make statistically significant in large samples but hard to make clinically meaningful. Reporting the proportion of patients who reach a patient-anchored acceptable state, and the absolute difference in that proportion between groups, restores clinical sense to the analysis and aligns the statistic with the question patients actually ask: "Am I going to be all right?" When ceiling effects are present,

a PASS-based outcome alongside the difference in means will increase the clinical impact of the research.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

Funding

The author(s) declared that financial support was not received for this work and/or its publication.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declared that generative AI was used in the creation of this manuscript. The authors used AI to optimize grammar and spelling of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Rosenberg J, Gram-Hanssen A, Reistrup H, Baker JJ. Redefining success in hernia surgery: the case for patient-reported outcomes. *J Clin Med* (2025) 14:6131. doi: 10.3390/jcm14176131

2. Wang Y, Devji T, Carrasco-Labra A, King MT, Terluin B, Terwee CB, et al. A step-by-step approach for selecting an optimal minimal important difference. *BMJ* (2023) 381: e073822. doi: 10.1136/bmj-2022-073822

3. Ahmed U, Baker JJ, Rosenberg J, Reistrup H. Turning a nation into a cohort: combining national registers and a digital mailbox for population-based health research. *Eur J Epidemiol* (2026), in press. doi: 10.1007/s10654-026-01408-z

4. Tubach F, Ravaud P, Baron G, Falissard B, Logeart I, Bellamy N, et al. Evaluation of clinically relevant states in patient reported outcomes in knee and hip osteoarthritis: the patient acceptable symptom state. *Ann Rheum Dis* (2005) 64:34–7. doi: 10.1136/ard.2004.023028

5. Tubach F, Dougados M, Falissard B, Baron G, Logeart I, Ravaud P. Feeling good rather than feeling better matters more to patients. *Arthritis Rheum* (2006) 55:526–30. doi: 10.1002/art.22110

6. Finley CL, Gray CF, Parvataneni HK, Prieto HA, Miley EN. Evaluating measures of clinical meaningfulness for patient-reported outcome measures in total joint arthroplasty: a systematic review. *J Arthroplasty* (2026) S0883-5403(26), 3162–75. doi: 10.1016/j.arth.2026.01.049

7. Gram-Hanssen A, Baker JJ, Reistrup H, Andersen KK, Rosenberg J. Protocol for the AFTERHERNIA project: patient-reported outcomes of groin and ventral hernia repair. *Hernia* (2025) 29:79. doi: 10.1007/s10029-025-03259-1

8. Mauch JT, Enriquez FA, Shea JA, Barg FK, Rhemtulla IA, Broach RB, et al. The abdominal Hernia-Q: development, psychometric evaluation, and prospective testing. *Ann Surg* (2020) 271:949–57. doi: 10.1097/SLA.0000000000003144

9. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. *Stat Med* (2006) 25:127–41. doi: 10.1002/sim.2331

10. Khan S, Zemberi J, Olsen T, McCutchen A, Voytik M, Jarrar S, et al. Determining the minimal important change of the abdominal Hernia-Q. *Hernia* (2026) 30:202. doi: 10.1007/s10029-026-03645-3