http://dx.doi.org/10.2106/JBJS.21.00773

October 28, 2021

What is the Minimum Clinical Important Difference and is it Helpful? Letter to the Editor

Daniel L. Riddle
Professor
Departments of Physical Therapy, Orthopaedic Surgery and Rheumatology, Virginia Commonwealth University, Richmond, VA, USA

Other Contributors:

Levent Dumenci
Professor
Department of Epidemiology and Biostatistics, 1301 Cecil B. Moore, Ave., Ritter Annex, Room 939, Temple University, Philadelphia, PA, USA 19122

Karhade and colleagues wrote on an important topic in healthcare, application and interpretation of the minimal clinically important difference (MCID) (1). Given the current heavy emphasis on patient-reported outcomes (PROs), interest is high for interpreting the meaning of changes in these measures. The MCID family of change measures is one of the more common ways of interpreting PRO change scores. Our reason for writing this letter is to highlight the irony of what we believe was the authors' recommendations – despite many problems with MCID-based investigations and the substantial conceptual problems with MCID measurement; surgeons should continue to rely on MCID to interpret the clinical relevance of PRO changes in their patients. We believe this recommendation is misguided.

The paper by Karhade et al (1) clearly defined several methodological inconsistencies in the MCID literature. These included but were not limited to conceptually distinct methods for calculating MCID via either anchor-based or distribution-based methods, highly variable and arbitrarily defined MCID thresholds, varying definitions of timeframes of interest, variations in types of external anchors and variations in distribution based methods and methods for dealing with loss to follow-up. Finally, the authors emphasized extensive variation in MCID estimates for patients with the same condition. We noted a similarly problem with extensive variation (2) using the example from te Molder et al (3) who reported 47 different definitions of good versus poor PROs for knee arthroplasty, most of which were based on the MCID family of measures.

The problems with MCID do not end with study execution. The original definition of MCID first used by

Jaeschke et al (4) was “The smallest difference in score in the domain of interest which patients perceive as beneficial and which would mandate, in the absence of troublesome side effects and excessive cost, a change in the patient’s management.” We suspect most would agree on two things. First, the definition of MCID by Jaeschke et al describes a clinically important construct and second, current methods do not capture this construct because costs, side effects or management are not considered.

Anchor based methods for MCID also lack conceptual clarity. Norman and colleagues found that single item retrospective ratings of change are typically highly correlated with the current state but uncorrelated or only weakly correlated with the baseline state (5). These data suggested that patients are not able to accurately recall their initial state, and accurately subtract this initial state from their follow-up state to derive the change. In other words, patients’ global ratings are driven largely by how patients are doing at follow-up and not by their baseline state. An additional noteworthy conceptual problem with MCID methods is the use of single item retrospective ratings to judge clinical relevance of multi-item PROs. PROMIS scales, for example, are highly sophisticated multi-item instruments developed via rigorous analytic and psychometric methods (6). Why would we rely on a crude and biased single item global rating scale to interpret them?

In our view, asking “what is the right threshold” for judging clinical important change is asking the wrong question. We suggested an alternative to the MCID family of change measures (2) when interpreting outcome using PROs. We are not the first to do so. Over three decades ago, Rindskopf et al, proposed latent class analysis as an alternative to the use of arbitrary threshold scores (7) (such as those used to calculate the MCID). We further documented conceptual problems for seeking “objective thresholds” for continuous scores like those from PROMIS instruments (8). We also argued that continued use of the MCID family of change score methods would lead to even more variability in estimates of important change and no progress in the interpretation of change scores. Much like Rindskopf et al, we argued for a statistically grounded latent class modeling method of defining different outcome trajectories (2,9). We have since demonstrated that this method can produce stable estimates of good versus poor outcome, even when outcomes of patients from different countries are compared (10). The benefits of a statistical model-based approach combined with evidence of stability across patient samples argue for an alternative to the substantial problems with MCID highlighted in our letter and in the paper by Karhade and colleagues (1).

References

References

1. Karhade A, Bono C, Schwab J, Tobert D. Minimum Clinically Important Difference: A Metric That Matters in the Age of Patient-Reported

Outcomes – PubMed. J Bone Jt Surg Am. 2021;(online ahead of print).
2. Riddle DL, Dumenci L. Classifications of good versus poor outcome following knee arthroplasty should not be defined using arbitrary criteria. BMC Musculoskelet Disord. 2020;21(1):604.
3. te Molder ME, Smolders JM, Hesterbeek PJ, Van Den Ende CH. Definitions of poor outcome after total knee arthroplasty: an inventory review. BMC Musculoskelet Disord. 2020;21:378.
4. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. Control ClinTrials. 1989;10:407-415.
5. Norman G, Stratford P, Regehr G. Methodological problems in the retrospective computation of responsiveness to change: the lesson of Cronbach. J Clin Epidemiol. 1997;50(8):869-879.
6. Fries J, Cella D, Rose M, Krishnan E, Bruce B. Progress in assessing physical function in arthritis: PROMIS short forms and computerized adaptive testing. J Rheumatol. 2009;36(9):2061-2066.
7. Rindskopf D, Rindskopf W. The value of latent class analysis in medical diagnosis. Stat Med. 1986;5(1):21-27.
8. Dumenci L. The Psychometric Latent Agreement Model (PLAM) for Discrete Latent Variables Measured by Multiple Items. Organ Res Methods. 2011;14(1):91-115.
9. Dumenci L, Perera R, Keefe F, Ang D, Slover J, Jensen M, et al. Model-based pain and function outcome trajectory types for patients undergoing knee arthroplasty: A secondary analysis from a randomized clinical trial. Osteoarthr Cartil. 2019;27(6):878-884.
10. Riddle DL, Macfarlane GJ, Hamilton DF, Beasley M, Dumenci L. Cross-Validation of Good versus Poor Self-reported Outcome Trajectory Types Following Knee Arthroplasty. Osteoarthr Cartil. 2021:s1063.

Conflict of Interest: None Declared

Article Author Response

1 November 2021

Article Author(s) to Letter Writer(s)

Drs. Riddle and Dumenci,

Thank you for your letter addressing our review article on minimum clinically important difference (MCID). Many of the points in your letter underscore our rationale for writing the manuscript. In particular, we appreciate your concern for anchor-based MCID calculation methods. This burgeoning approach to MCID in the literature belies confounding from recall bias and an inherent statistical misstep.

Yet, substitution with latent class analyses (LCA) is not a panacea for the faults of MCID methods. LCA is less reliable with small sample sizes, or if a class size becomes small. This is problematic if applied to total hip arthroplasty, for example, with historical success rates above 95%. Furthermore, it is pragmatic for clinicians, researchers and policymakers to judge clinical improvement in a binary fashion (yes/no). If a three- or four-class model is a more appropriate fit for the data than a two-class model, the interpretation becomes ambiguous. Finally, a criticism of MCID as arbitrary holds its own irony juxtaposed to the nominal fallacy associated with the “good” and “poor” labels often used in latent class analysis.

http://dx.doi.org/10.2106/JBJS.21.00773

Despite its own shortcomings, LCA is a powerful statistical technique that has advantages in certain scenarios. I see the different statistical approaches as synergistic as the research community searches for a reliable definition of clinical improvement. This paradigm accesses the strengths of the individual methodologies—and in combination (1). Rather than misguided, this manuscript provides a practical approach for readers if MCID is the determinant of significance for a research study.

Daniel G. Tobert MD

Massachusetts General Hospital

Boston, MA

References

1. Yuksel S, Ayhan S, Nabiyev V, et al. Minimum clinically important difference of the health-related quality of life scales in adult spinal deformity calculated by latent class analysis: is it appropriate to use the same values for surgical and nonsurgical patients? Spine J. 2019;19(1):71-78. doi:10.1016/j.spinee.2018.07.005