Title

Minimally important changes do not always reflect minimally important change; moreover, there is no need for them.

Journal

Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation, 32(5)

ISSN

0962-9343

Authors

Peipert, John Devin
Cella, David
Hays, Ron D

Publication Date

2023-05-01

DOI

10.1007/s11136-023-03366-5

Peer reviewed

Quality of Life Research (2023) 32:1403–1404
https://doi.org/10.1007/s11136-023-03366-5

LETTER TO THE EDITOR

Minimally important changes do not always reflect minimally important change; moreover, there is no need for them

John Devin Peipert[1] · David Cella[1] · Ron D. Hays[2]

Accepted: 2 February 2023 / Published online: 13 February 2023

In response to the recently published letter: "Likely change indexes do not always index likely change; moreover, there is no need for them" [1], we suggest instead that this sentiment applies to "minimally important change." The author correctly points out that the likely change index (LCI) [2] is similar to the reliable change index (RCI), a significance test for within individual change. The RCI is calculated as: $(X_1 - X_2)/\sqrt{2}SEM$, where the X_1 and X_2 are PRO scores for an individual patient at baseline and a follow-up, the SEM (standard error of measurement) is the $SD_1\sqrt{1 - reliability}$ (SD_1 = standard deviation at baseline, we note that the SD of the change score can be used instead). The LCI was proposed because many patients who report that they have changed (gotten better or worse) will be classified as not changed using the conventional $p < 0.05$ threshold with the RCI. The amount of change (critical value) needed to be significant on the RCI is known as the coefficient of repeatability: $criticalvalue * \sqrt{2}SEM$. For $p < 0.05$, the critical value is 1.96, but this value decreases as the p-value increases. As we demonstrate in our article, LCI thresholds for 68% and 50% confidence tended to align more with anchor-based estimates of meaningful change and, we suggest, may produce more accurate individual classification overall [2]. The first assertion in the letter is that LCI interpretation relies upon a statistical fallacy: that the significance level used for the LCI will indicate the probability of real change–for example, that the LCI using a p-value of $p = 0.32$ would indicate at 68% probability that the patient experienced true change. While it is correct that this interpretation would support a statistical fallacy, we did not use this interpretation for the LCI. We also do not agree that using other p-values than $p < 0.05$ is "awkward." It is increasingly recognized in scientific and statistical communities that the dogma of adhering to $p < 0.05$ can be more harmful than beneficial.

We call further attention to a series of simulation analyses offered in the letter suggesting the LCI does not indicate a high likelihood of true change. Analyses presented in the letter actually show the opposite, providing support for the LCI. According to those simulations, when the prevalence of true change is 50% and reliability of a PRO is high (0.90), observed change scores at the level of LCIs at 50% ($p = 0.50$), 68% ($p = 0.32$), and 95% ($p = 0.05$) confidence imply true change probabilities of 0.80, 0.87, and 0.97, respectively. When the prevalence of true change is 40%, the probabilities of observing true change with LCI's at 50%, 68%, and 95% are 0.62, 0.73, and 0.92 respectively. Only when the PRO's reliability is decreased to 0.70 do probabilities of true change drop to unacceptable levels (0.38, 0.48, and 0.77, respectively). For many applications, a change threshold that imparts > 70% probability that the patient has changed will be deemed sufficient and attractive to researchers if it is close to the amount of change patients, on average, find meaningful. Under scenarios with low PRO reliability and low prevalence of true change, the LCI would lead to misclassifying patients who have not changed as "changed." This should not come as a surprise. General guidance around PRO reliability has cautioned that high reliability is required to interpret scores for individuals [3]. PROs with low reliability should not be used to measure change in individuals. This advice applies to using the LCI and any other method for individual patients.

We take this opportunity to point out another case where the LCI's goal has been misconstrued. The author states that the typical application for the LCI would be with an individual patient in a clinical setting; moreover, the author noted that in research, there is no need to pinpoint whether individual patients have changed. This is false. Although the

✉ John Devin Peipert
john.peipert@northwestern.edu

[1] Department of Medical Social Sciences, Northwestern University Feinberg School of Medicine, 625 Michigan Ave, 21st Floor, Chicago, IL 60611, USA

[2] Division of General Internal Medicine and Health Services Research, Department of Medicine, University of California Los Angeles, Los Angeles, CA, USA

LCI may have useful clinical applications, the original idea that led to the LCI was aimed at clinical research applications like time to deterioration analyses, wherein knowing whether an individual has changed is required. Using the LCI allows the researcher to know the probability of classifying individual patients as having changed (in this case, deteriorated) by chance alone, which in turn gives important information about how much the PRO's measurement error affects such classification. The LCI could improve time to deterioration analyses, since the standard approach is to use group-based, anchor methods [4], which leads to misclassification of individual patients [5]. We caution that indiscriminately applying statistics like the anchor-based, minimally important change (MIC) statistic without regard for measurement error would often lead to classifying individuals as having changed when they have not. Of course, it is unwise to apply any statistic indiscriminately, including the LCI. For this reason, we noted in our paper that "the level of confidence used with the LCI should reflect the needs of the application," while cautioning "potential users that as the confidence level approaches 50%, the likelihood of that change being due to chance increases" [2]. We do indeed agree with the author that using other sources of information about change, especially in clinical settings, is a good idea, and that the LCI can be paired with other sources of information as appropriate. In that regard, one could argue that an MIC does not always reflect an MIC, and moreover, there is no need for it.

References

1. Terluin, B. (2022). Likely change indexes do not always index likely change; moreover, there is no need for them. *Quality of Life Research*. https://doi.org/10.1007/s11136-022-03314-9
2. Peipert, J. D., Hays, R. D., & Cella, D. (2022). Likely change indexes improve estimates of individual change on patient-reported outcomes. *Quality of Life Research*. https://doi.org/10.1007/s11136-022-03200-4
3. Lord, F. M. (1956). The measurement of growth. *Educational and Psychological Measurement, 16*(4), 421–437. https://doi.org/10.1177/001316445601600401
4. Charton, E., Cuer, B., Cottone, F., et al. (2020). Time to deterioration in cancer randomized clinical trials for patient-reported outcomes data: a systematic review. *Quality of Life Research, 29*(4), 867–878. https://doi.org/10.1007/s11136-019-02367-7
5. Hays, R. D., & Peipert, J. D. (2021). Between-group minimally important change versus individual treatment responders. *Quality of Life Research, 30*(10), 2765–2772. https://doi.org/10.1007/s11136-021-02897-z