Contents lists available at ScienceDirect

Journal of Hand Therapy

journal homepage: www.elsevier.com/locate/jht

Letter to the Editor

Comments on "Concurrent validity and precision of the thumb disability examination (TDX) in first carpometacarpal osteoarthritis" by Johnson et al.

I read the article with great interest by Johnson et al. entitled "Concurrent validity and precision of the thumb disability examination (TDX) in first carpometacarpal osteoarthritis."[1] The authors of the study were purposed to reveal the concurrent validity, test-retest reliability, standard error of measurement (SEM) and minimal detectable change (MDC) of the TDX. I congratulate the authors of the study; as they also stated in their article, there is no available tool for assessing osteoarthritis in the thumb carpometacarpal joint. While I believe the publication provides an essential contribution to the literature, there are some methodological concerns that I would like to address which may affect the results of the study. In this way, we can exchange ideas with the authors to contribute to the methodology of further studies.

Firstly, it is crucial to reconsider the sample size of the study (n: 16). In the last part of the method section, the authors confirmed that the required number for the sample size was 14. In their calculations, at 80% power and 0.05 statistical significance level, the calculation was carried out with an effect size based on a "Spearman correlation coefficient" of 0.70, which belong to a similar reference article. However, this value did not provide sufficient data for calculating the intraclass correlation coefficient (ICC) based on test and retest in the current study. Because correlational analysis (r) is based on a single measurement, reliability is based on two measurements. In this respect, it is necessary to calculate the study's sample size separately for the retest. For this, the formula for calculating the minimum number required for ICC proposed by Walter and Donner should be considered.[2] A sample calculation based on the current study data reveals that at least 43 individuals are required in the repeat measurement for the minimum acceptable ICC (0.60), the expected ICC (0.81: the value obtained in this study), the significance level of 0.05, and the power of 0.80. The authors' resignation of additional sample size or power analysis of the retest measurement will further confirm the evidence of the study.

Secondly, another common recommendation for the sample size calculation is based on Terwee et al. Ten times the number of survey items was suggested in the psychometric analysis studies.[3] TDX have 20 questions so a total of 200 patients were required for this validity study. From this point of view, the current sample presents statistical data of very few patients. Data from a small number of patients may lead to a standard deviation that does not adequately represent the relevant population. This issue needs to be addressed in terms of the calculation of SEM and MDC. Because SEM was calculated based on the standard deviation value, and MDC is calculated with the SEM value.[4] In this respect, the current standard deviation may not represent the universe sample due to the low sample size. Accordingly, the MDC value may not provide a precise reference value in patient monitoring.

Thirdly, the test-retest reliability method should be reviewed. The authors performed two assessments with an interval of 7-21 days. They characterized this period as the most accepted evaluation period in hand therapy clinical practice. However, the issue investigated here is test-retest reliability and, according to common opinion, Terwee et al. reported that this period should be between 2 and 14 days.[3] It is also essential that patients should not receive any treatment during this period. In this respect, the authors should have excluded individuals utilizing any treatment agent during this period. The exclusion criteria also should have been expanded in this respect. The consistency between the two measurements may be masked by the potential treatment intervention, which may lead to a lower ICC value.

Last but not least, the type of ICC is not declared. It was necessary to decide on the test-retest reliability with the appropriate method in the Shrout and Fleiss classification and to perform the analysis accordingly.[5] I would welcome the comments of the authors to address these issues, especially since the providing of reliability of the inertial sensors will further validate the findings of the study.

References

1. Johnson L, Karau R, McGee C. Concurrent validity and precision of the thumb disability examination (TDX) in first carpometacarpal osteoarthritis. *J Hand Ther.* 2021;35:428–434. https://doi.org/10.1016/j.jht.2021.07.002
2. Walter SD, Eliasziw M, Donner A. Sample size and optimal designs for reliability studies. *Stat Med.* 1998;17:101–110. https://doi.org/10.1002/(sici)1097-0258(19980115)17:1<101::aid-sim727>3.0.co;2-e
3. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol.* 2007;60:34–42. https://doi.org/10.1016/j.jclinepi.2006.03.012
4. Portney LG, Watkins MP. *Foundations of Clinical Research: Applications to Practice.* Upper Saddle River, NJ: Pearson/Prentice Hall; 2009.
5. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull.* 1979;86:420–428.

Fatih Özden, PT, MSc
Muğla Sıtkı Koçman University, Köyceğiz Vocational School of Health Services, Department of Health Care Services, Division of Elderly Care, Köyceğiz, Muğla, Turkey
E-mail address: fatihozden@mu.edu.tr (F. Özden)