

Use of Ordinal Measures in JHTR Manuscripts

Limitations of Ordinal Measurement

Many measures and scales used in rehabilitation are ordinal measures. The score on an ordinal measure usually represents the simple sum of the score on each item. Although a higher score on the measure indicates that the person being evaluated has more of whatever trait is being measured, the score does not indicate how much more. For example, a higher score on the Disability Rating Scale (DRS) indicates greater overall disability. However, judging from the percentage of individuals with brain injury who are not disabled on various items at one-year follow-up, independence in basic self-care is markedly less challenging than independent living which, in turn, is less challenging than employability. That is, lack of independence in self-care appears to indicate much more disability than limitations in living independence. Because the measure is ordinal, how much more is unknown. Some measures like the Glasgow Outcome Scale-Extended (GOS-E) are single-item measures that describe a series of ordered categories of recovery. As with other ordinal measures, the degree of recovery from one level to the next cannot be known; only that a higher level indicates greater recovery.

Converting Ordinal Measures to Interval Scales

Precision. The limited precision of ordinal measures has been known for many years in rehabilitation^{1,2} and has been addressed by converting existing measures or designing new measures that provide an interval scale through Rasch or other item-response theory (IRT) methods.^{3,4} A higher score on an interval measure indicates not only that the person being evaluated has more of a trait but how much more. For instance, the distance between 3 inches and 5 inches on a ruler is exactly the same as the distance between 5 and 7 inches. If a Rasch metric (not raw scores) is used to report results of a FIM assessment, the difference in overall required assistance between scores of 20 and 25 would be the same as between those with scores of 25 and 30. In the last 30 years, a number of interval-level rehabilitation scales have been developed or converted from existing scales using Rasch or other IRT methods offering our field measures with greater precision.

Coverage. Rasch and IRT methods also attend to coverage, i.e., assuring that the trait is measured across its entire spectrum. Many ordinal measures in rehabilitation do not demonstrate adequate coverage and show dramatic floor (a large percentage of very low scores like the DRS) or ceiling (large percentage of very high scores like the GOS-E) during the post-acute period.

When ordinal measures may be acceptable

While the editors of JHTR encourage the use of modern interval measures, this is not a requirement for publication. In some cases, the use of an ordinal measure may be acceptable. For instance, if the outcome of interest is whether an individual crosses a specified threshold of recovery, e.g., independence in basic self-care, an ordinal measure is probably sufficient. Similarly, if specific levels of outcome are of interest, for example, the level of depressive symptoms, an ordinal measure like the PHQ-9 would be acceptable. However, if the question is how depressed the participants feels, an interval-level measure of depressed mood would be preferred. In large database analyses, only summary raw scores for ordinal measures may be available. However, in this latter case, the use of an ordinal measure should be listed as a limitation with the suggestion that future research use a measure with greater precision.

Statistics

Nonparametric statistics should be used for ordinal measures (e.g., Mann-Whitney, Kruskal-Wallis, Spearman correlation); whereas, parametric statistics (e.g., t-test, ANOVA, Pearson correlation) can be used with interval measures. However, nonparametric methods are not available for all statistical operations, such as, linear regression. There are regression methods for ordered categories, but these may not be reasonable for all ordinal measures, especially those with a large number of categories. If parametric statistics are used with ordinal measures because of the unavailability of an analogous nonparametric procedure, this should be listed as a limitation.

Guidance for Reviewers

As mentioned above, the use of ordinal measures is not prohibited for publication in JHTR. Reviewers should use their best judgement in commenting on measures. Examples of possible reviewer comments follow:

Precision: If an ordinal measure is being used in a study, particularly as a primary outcome measure, a reviewer might suggest: *“The use of an ordinal scale as an outcome measure limits measurement precision and should be listed as a limitation. Authors should also consider suggesting that future research use an interval-level measure to improve measurement precision.”*

Coverage: *“The measure appears to show a significant ceiling (or floor) effect. This should be listed as a limitation with the suggestion that future research use a measure with greater precision and sensitivity to those at high (or low) levels.”*

Statistics: *“Because ordinal measures are being used, analyses should be recalculated using nonparametric methods.”*

“While nonparametric methods for linear regression are not available, if methods for ordered categories are feasible and would provide interpretable results, they should be considered. If not, the use of an ordinal measure as the dependent variable should be listed as a limitation.”

References

1. Merbitz C, Morris J, Grip JC. Ordinal scales and the foundations of misinference. Arch Phys Med Rehabil 1989;70(4):308-12. PMID: 2535599
2. Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil. 1989 Nov;70(12):857-60. PMID: 2818162
3. Bond T, Yan Z, Heene M. Applying the Rasch Model: Fundamental Measurement in the Human Sciences(4th ed). Routledge, 2021.
4. Velozo CA, Seel RT, Magasi S, Heinemann AW, Romero S. Improving measurement methods in rehabilitation: core concepts and recommendations for scale development. Arch Phys Med Rehabil 2012;93(8 Suppl 2):S154-63. PMID: **22840881** DOI: [10.1016/j.apmr.2012.06.001](https://doi.org/10.1016/j.apmr.2012.06.001)