What does Cronbach's Alpha tell us about the EQ-5D? A methodological commentary to "Psychometric properties of the EuroQol Five Dimensional Questionnaire (EQ-5D-3L) in caregivers of autistic children"

In a paper just accepted for publication in Quality of Life Research, Khanna et al. [1] investigate the psychometric properties of the EQ-5D with three levels [2] using data from a sample of 316 US—American caregivers of autistic children. The authors examine convergent and discriminant validity of the five EQ-5D dimensions by analyzing the relation of these dimensions with the dimensions of the SF-12v2 [3], and they examine known group validity by analyzing how the severity of the children's autism relates to the EQ-5D index and the VAS scale. With a few exceptions, the results of these examinations accord with the authors' expectations. The authors interpret this as evidence for the EQ-5D's validity. The authors also examine the reliability of the EQ-5D. For this purpose, they compute Cronbach's alpha and item-total correlations for the five dimensions of the EQ-5D. Cronbach's alpha is only 0.63, whereas the item-total correlations range from 0.25 for anxiety/depression to 0.58 for usual activities. Following the recommendations of Nunnally [4], the authors judge the result for Cronbach's alpha as unsatisfactory and the results for the item-total correlations as acceptable.

Whilst the authors' investigations concerning validity are perfectly reasonable, those concerning reliability warrant some critical remarks. The method they apply, that is, the estimation of reliability via Cronbach's alpha and item-total correlation, is not adequate for the kind of measurement instrument investigated, that is, for an index measurement instrument like the EQ-5D.

The approach of estimating the reliability of a measurement instrument via Cronbach's alpha and item-total correlations has clear limitations. It is not adequate for all kinds of measurement instruments, but only for homogenous tests. In a perfectly homogenous test, all items reflect the same dimension, and therefore, deviations from perfect inter-item correlations reflect nothing else

but measurement error. For such a test, the deviation of Cronbach's alpha from the upper bound 1 is a direct function of measurement error, and consequently, Cronbach's alpha itself is a valid estimation of the test's reliability. The opposite of homogenous tests are heterogeneous tests. A test is heterogeneous when it is composed of items which address qualitatively different aspects or even qualitatively different dimensions of the object of investigation. In a heterogeneous test, the different items cannot correlate perfectly even if there is no measurement error. Consequently, when Cronbach's alpha is applied to heterogeneous tests, the true reliability is underestimated. The more heterogeneous a test is, the more the reliability will be underestimated. Cronbach himself already explicitly pointed out this feature of Cronbach's alpha when he first presented this coefficient [5]. Item-total correlations have similar features like Cronbach's alpha. The more heterogeneous the test is, the lower the item-total correlations will be.

A good index measurement instrument requires a set of questions which cover all relevant different aspects of the object of investigation. Ideally, the different questions should reflect different analytically independent dimensions, because in this case, the informative value provided by each single question would be maximal. Selecting items with regard to this ideal necessarily produces a very heterogeneous test. Consequently, neither Cronbach's alpha nor item-total correlations for such a test can be particularly high. It would even be extremely unwise to strive for a high Cronbach's alpha and high item-total correlations, because selecting items with regard to maximizing these statistics would in fact impede the construction of a measurement instrument which covers all relevant aspects of the object of investigation. The items of the EQ-5D have actually been selected to cover all the different aspects of health which were considered, at the time of item selection, as relevant for health-related quality of life. Consequently, a high Cronbach's alpha can neither be expected nor desired for this set of items. The same holds for the item-total correlations.

To sum up, both Cronbach's alpha and item-total correlations decrease with measurement error and with heterogeneity. For a good index measurement instrument, the first should be minimized, but the second should be maximized. Therefore, Cronbach's alpha and the item-total correlations tell us

nothing about the quality of an index measurement instrument. They should not be applied as criteria for quality. If information about reliability is desired, the test–retest approach should be applied.

References

1. Khanna, R., Jariwala, K., Bentley, J. P. (2013). Psychometric properties of the EuroQol Five Dimensional Questionnaire (EQ-5D-3L) in caregivers of autistic children. Quality of Life Research [Epub ahead of print].

2. The EuroQol Group. (1990). EuroQol—a new facility for the measurement of health-related quality of life. Health Policy, 16(3), 199–208.

3. Ware, J. E., Kosinski, M., Turner-Bowker, D. M., & Gandek, B. (2002). User's manual for the SF-12v2 health survey (with a supplement documenting SF-12 health survey). Lincoln, RI: Quality Metric Incorporated.

4. Nunnally, J. C. (1978). Assessment of reliability. Psychometric theory (2nd ed., pp. 225–255). New York: MacGraw-Hill Series.

5. Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297–334.