available at www.sciencedirect.com

SCIENCE DIRECT®

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

CORRESPONDENCE

The Chimera of Population Norms

It was with great interest that we read the recent report by Gilet and colleagues in the recent edition of Value in Health [1]. Gilet et al. [1] attempted to develop a normative database for the quality of life assessment in growth hormone deficient adults (QoL-AGHDA). Unfortunately, the article demonstrates a number of misunderstandings of the nature of patient-reported outcome measurement, the value of population norms, and the application of Rasch analysis.

The QoL-AGHDA is a measure of quality of life (not health-related quality of life as stated in the article) that is specific to patients with growth hormone deficiency. Its content was derived from such patients and was never intended to be completed by healthy individuals. For the QoL-AGHDA to be valid for use with a healthy population it would first be necessary to show that it is psychometrically strong for such a group and that their responses to the items fit the Rasch model. In order to allow comparisons to be made between growth hormone deficient (GHD) patients and a general population it would also be necessary to establish that the scale works in the same way with both groups. None of these issues was adequately addressed in the article.

The authors do not indicate whether or not their data fit the Rasch model – a major omission. If this is not the case the item order reported in Figure 2 might be invalid. The authors also state that there are differences between the scores of males and females without reporting whether these differences result from differential item functioning (DIF). The authors used the software program RUMM for the Rasch analysis. This program provides different types of evidence for unidimensionality, including overall fit statistics and tests for DIF [2].

It is clear from the Person-Item map presented that the QoL-AGHDA, as expected, is totally unsuitable for use with a healthy population. Approximately half of the healthy population sample has less impairment than is measured by the items in the questionnaire, which accounts for the large basement effect observed. The map also reveals that the scale has marked redundancy for this sample. The highly biased scores would prevent reference values based on means being produced. In the absence of a comparison of scaling properties between their sample and a sample of GHD patients, no conclusions can be drawn about the possibility of establishing normative values.

The concept of producing norms in this way is also of debatable scientific merit. It is an attempt to compare apples with pears. Why should healthy individuals fit on a scale of the "impact of GHD on the quality of life of adults"? Similarly, why should they fit on the same scale as patients with other chronic diseases as recommended by the authors? This problem applies equally to the application of any measure – including generic health status measures such as the Short Form 36 health survey, Nottingham Health Profile, or EuroQol five-dimensional (EQ-5D) questionnaire. Such measures cannot be used to make valid comparisons between healthy and diseased populations or between groups of patients with different illnesses. The whole concept of normative values is also questionable in scaling terms due to the existence of DIF between the multiple sub-populations in the reference sample.

However, Rasch analysis does provide a means for making valid comparisons between the impacts of different diseases. This involves making disease-specific measures based on the same measurement model and applying co-calibration of scores on the different scales. Interestingly, the QoL-AGHDA uses the needs-based model of quality of life [3,4], which is the only widely operationalized model in the field of patient-reported outcome measurement [5,6]. The needs-based model has been employed in the development of over 30 disease-specific quality of life measures. Such a body of measures begins to provide a means for making valid comparisons between diseases. Furthermore, by using a similar methodology, valid comparable utility values can be generated specific to different diseases. In this way the utilities derived from the quality of life scales are able to avoid the weaknesses of the generic utility measures such as the EQ-5D that have limited reproducibility and responsiveness.

Stephen P. McKenna, PhD
Galen Research, Manchester, United Kingdom

John Brodersen, MD
University of Copenhagen, Copenhagen, Denmark

1098-3015/$36.00 – see front matter

Published by Elsevier Inc.
doi:10.1016/j.jval.2010.11.019

REFERENCES

[1] Gilet H, Chachuat A, Viala-Danten M, et al. Application of the disease-specific quality of life assessment of hormone deficiency in adults (QoL-AGHDA) questionnaire in a general population: results from a French panel study. Value Health 2010;13:495–500.
[2] Brodersen J, Meads DM, Kreiner S, et al. Methodological aspects of differential item functioning in the Rasch model. J Med Econ 2007;10:309–24.
[3] Hunt SM, McKenna SP. The QLDS: a scale for the measurement of quality of life in depression. Health Policy 1992;22:307–19.
[4] McKenna SP, Doward LC, Alonso J, et al. The QoL-AGHDA: an instrument for the assessment of quality of life in adults with growth hormone deficiency. Qual Life Res 1999;8:373–83.
[5] McKenna SP, Doward LC. The needs-based approach to quality of life assessment. Value Health 2004;7(Suppl. 1):S1–3.
[6] McKenna SP, Doward LC, Meads DM, et al. Summary of needs-based quality of life instruments. Value Health 2004;79(Suppl. 1):S39–40.