RASCH ANALYSIS OF THE PEDIATRIC QUALITY OF LIFE INVENTORY 4.0 GENERIC CORE SCALES ADMINISTERED TO PATIENTS WITH DUCHENNE MUSCULAR DYSTROPHY

P.345

Erik Landfeldt, PhD[1,2], Joel Iff, PhD[3], Erik Henricson, PhD, MPH[4], and the CINRG DNHS Investigators

[1]Department of Women's and Children's Health, Karolinska Institutet, Stockholm, Sweden; [2]ICON plc, Stockholm, Sweden; [3]Sarepta Therapeutics, Inc, Cambridge, Massachusetts, USA;
[4]Department of Physical Medicine & Rehabilitation, University of California Davis School of Medicine, Sacramento, CA, USA.

INTRODUCTION

Duchenne muscular dystrophy (DMD) is a rare, X-linked, severely debilitating, and ultimately fatal neuromuscular disease characterised by progressive muscle weakness.[1] The discovery of several novel treatment strategies for DMD has resulted in a need to identify patient-reported outcome scales that are fit for purpose to quantify drug benefits, to inform regulatory approval and reimbursement.[2] One of the most commonly applied measures of health-related quality of life (HRQoL) in paediatric populations, including DMD, is the Pediatric Quality of Life Inventory 4.0 Generic Core Scales (PedsQL 4.0 GCS).[3]

OBJECTIVE

The objective of this study was to explore the psychometric properties of the PedsQL 4.0 GCS administered to patients with DMD using Rasch analysis.[2]

METHODS

PATIENTS

Patients with DMD were recruited from 20 centres across nine countries as part of the CINRG Duchenne Natural History Study (DNHS).[4-6] In the DNHS, patients were asked to complete the PedsQL 4.0 CGS with the help of a caregiver, if needed.

STATISTICAL ANALYSIS

Rasch analysis is the formal testing of a scale against a mathematical model.[7] A Rasch partial credit model was fitted,[8] determined based on a likelihood-ratio test, to the PedsQL 4.0 GCS data using RUMM2030. Individual item misfit was defined as a fit residual >|2.5| or a χ^2 Bonferroni-adjusted p-value <0.002174 (0.05/23).[9,10] Other analyses included; person fit to the Rasch model (defined as a fit residual >|2.5|); ordering of item response category thresholds (i.e., that respondents are able to differentiate between response categories); local item dependency (i.e., if a reply to one item predicts the reply to another item); targeting (i.e., the match of the different ability levels estimated through the Rasch model with the ability levels observed in the sample); reliability (Person Separation Index [PSI],[10,11] and Cronbach's α), differential item functioning (i.e., item stability) investigated through analysis of variance by disease stage and steroid use; and unidimensionality through principal components analysis of the residuals.[11]

RESULTS

ITEM FIT TO THE RASCH MODEL

In total, 329 patients with DMD (mean age: 9 years, range: 3–18 years; 75% ambulatory) completed the PedsQL 4.0 GCS. The distribution of replies is shown in Figure 1. The most difficult scale items, expressing the greatest loss in HRQoL, were those associated with emotional well-being (e.g., being teased by other children, feeling sad, and not making friends). The least difficult items were those mainly reflecting physical disability (e.g., lifting heavy objects, participating in sports, and running).

The mean item fit residual was estimated at 0.301 (SD: 1.385). The overall item-trait interaction χ^2 value was 178 (115 degrees of freedom, p<0.001), indicating that the items were not working as expected across different levels (i.e., class intervals) of HRQoL in the sample. Many items also exhibited non-trivial local dependency.

The PSI and Cronbach's α were estimated at 0.903 and 0.901, respectively. By investigating residual principal component loadings, the PedsQL 4.0 GCS was found to be multidimensional, with 24% (95% CI: 21%-26%) statistically significant t-tests (p<0.05).

Figure 1. Distribution of replies to the PedsQL 4.0 GCS

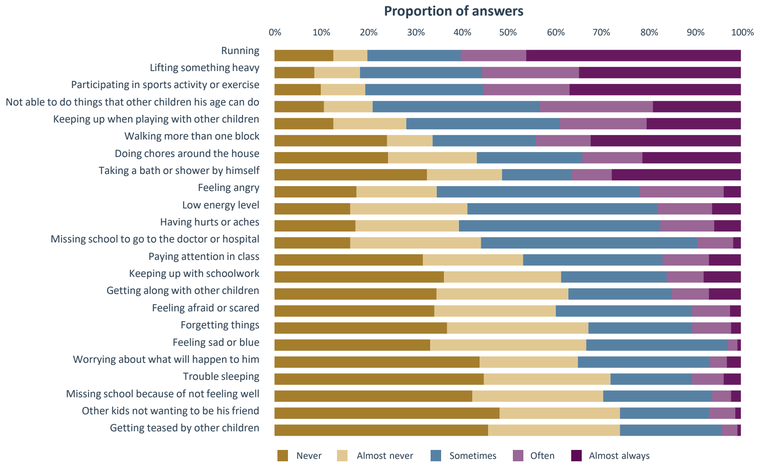
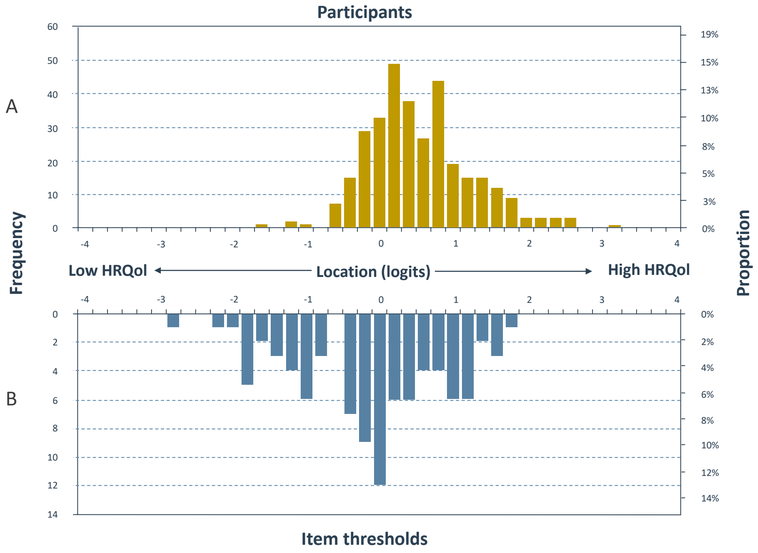


Figure 2. Distribution of participants (A) and item thresholds (B) on the estimated continuum



Panel A shows the location of the participants (n=329) on the interval logit scale representing level of HRQoL (a low number represents low HRQoL, and vice versa). **Panel B** shows the location of the PedsQL 4.0 GCS item thresholds (23×4=92) on the same logit scale (a low number represents low item difficulty, and vice versa).

ITEM THRESHOLDS

Disordered thresholds were identified for 87% (20 of 23) of all items, indicating that participants generally had difficulty discriminating between response categories given their level of HRQoL.

PERSON FIT TO THE RASCH MODEL

The mean location of individual responses was 0.453 (range: -0.854 to 0.992), indicating that the sample exhibited a higher level of HRQoL than what would be expected on average from the included items, with a mean fit residual of -0.255 (SD: 1.504) (Figure 2).

DIFFERENTIAL ITEM FUNCTIONING

The analysis also revealed significant issues with uniform differential item functioning by disease stage (early ambulatory vs. late ambulatory vs. non-ambulatory), but not glucocorticoid use (any lifetime exposure vs. no exposure).

DISCUSSION

Our Rasch analysis revealed several significant psychometric issues with the PedsQL 4.0 GCS as administered to patients with DMD, including disordered thresholds (suggesting that patients and caregivers perceive the current level structure of the scale as ambiguous), non-trivial local item dependency, suboptimal targeting, and poor item-trait interaction.

Interestingly, the most difficult PedsQL 4.0 GCS items were those associated with emotional well-being, as opposed to physical disability. These findings suggest that the largest loss in HRQoL was captured by items associated with morbidity and disability in addition to and beyond the primary somatic manifestation of DMD. However, in this context, it is important to bear in mind that many patients are likely to rely upon medical devices and aids to perform their daily functional tasks and activities as captured by the PedsQL 4.0 GCS, which would generate estimates of a relatively higher person ability and/or lower item difficulty for these items.

We found evidence of significant differential item functioning by disease stage, indicating that the scale does not behave similarly across the progression sequence in DMD.

Finally, our psychometric examination revealed that the PedsQL 4.0 GCS may not be regarded a unidimensional, interval rating-scale of HRQoL among patients with DMD. These results show that it is not meaningful to compare mean total scale scores across trials, studies, or samples, or even individual scores between patients, as they are not invariant.

CONCLUSION

The PedsQL 4.0 GCS may not be fit for purpose to measure HRQoL in patients with DMD and should be used with caution in this disease population until further evidence is made available.

ACKNOWLEDGEMENTS AND DISCLOSURES

This study was funded by Sarepta Therapeutics, Inc.; E. Landfeldt is employed by ICON plc., and received funding from Sarepta Therapeutics, Inc. for conduction the Rasch analysis, and writing support; J. Iff is an employee of Sarepta Therapeutics Inc. and may own stock/options in the company; E. Henricson, nothing to disclose.

REFERENCES

1. Emery. *Lancet* 2002; 359: 687–695. **2.** Hobart et al. *Lancet Neurol* 2007; 6: 1094–1105. **3.** Varni et al. *Med Care* 2001; 39(8): 800–812. . McDonald et al. *Muscle Nerve* 2013; 48(1): 32–54. **5.** McDonald et al. *Lancet* 2018; 391(10119): 451–461. **6.** Henricson et al. *Muscle Nerve* 2013; 48(1): 55–67. **7.** Rasch G. Probabilistic models for some intelligence and attainment tests (1st Edition). Copenhagen: Danish Institute for Education Research; 1960. **8.** Masters. *Psychometrika* 1982; 47: 149–174. **9.** Lundgren Nilsson et al. *J Rehabil* Med 2011; 43(10): 884–891. **10.** Tennant et al. *Arthritis Rheum* 2007; 57(8): 1358–1362. **11.** Smith EV. *J Appl Meas* 2002; 3(2): 205–231.

SAREPTA THERAPEUTICS