Psychometric Analysis of the Patient-Reported Plexiform Neurofibromas Quality of Life Measure Using KOMET Study Data

MSR190

Ayo Adeyemi,[1] Josephine Norquist,[2] Xiaoqin Yang,[2] Peter Annas,[3] Laurine Bunod,[4] Randolph de la Rosa Rodriguez,[1] Emma McMurtry,[5] Antoine Regnault[4]

[1]Alexion, AstraZeneca Rare Disease, Boston, MA, USA; [2]Merck & Co., Inc., Rahway, NJ, USA; [3]Alexion, AstraZeneca Rare Disease, Copenhagen, Denmark; [4]Modus Outcomes, a THREAD company, Lyon, France; [5]Clinical Pharmacology & Safety Sciences, AstraZeneca, Cambridge, UK

OBJECTIVES

- To evaluate the psychometric properties of PlexiQoL and to provide an estimate of clinically meaningful within-patient change in the context of the KOMET study data

CONCLUSIONS

- Overall, the quantitative evidence generated using data from KOMET confirmed the good measurement properties of the PlexiQoL total score
- A 2-point decrease is a tentative estimate of MSD for the PlexiQoL total score, based on distribution-based methods in the absence of a strong anchor variable

PLAIN LANGUAGE SUMMARY

Why did we perform this research?
This research aimed to show that a score obtained from a patient self-reported questionnaire called the plexiform neurofibromas quality of life questionnaire (PlexiQoL) allowed accurate measuring of the quality of life of adults with neurofibromatosis type 1 and plexiform neurofibromas (NF1-PN). Additionally, it aimed to determine how much change in the PlexiQoL score was perceived as meaningful for a patient in the context of the KOMET study data.

How did we perform this research?
The PlexiQoL questionnaire has 18 questions that cover five areas of life affected by NF1-PN. Data were collected from participants at multiple points during the study to assess their quality of life. Psychometric analyses were used to determine if the PlexiQoL provided reliable and meaningful measurements of quality of life.

What were the findings of this research and what are the implications?
The findings showed that the PlexiQoL tool was generally reliable and could effectively measure the quality of life of adult patients with NF1-PN. Most participants were able to complete the questionnaire, indicating its practicality. The distribution-based analysis suggested that a change of 2 points in the PlexiQoL score could be considered a meaningful change in a patient's quality of life. Additional studies are encouraged to further investigate these findings.

Where can I access more information?
More detailed information and the full results of this study can be accessed through the supplementary materials (available through the QR code).

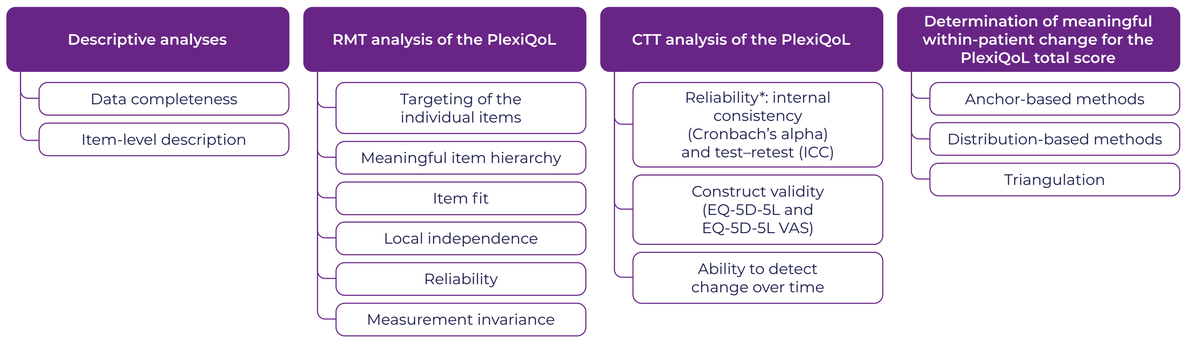
BACKGROUND

- Neurofibromatosis type 1 (NF1) is a rare, multisystem, genetic disorder that is caused by mutations in the *NF1* tumor suppressor gene and is associated with a diverse range of clinical manifestations;[1–4] up to 50% of patients with NF1 develop nerve sheath tumors called plexiform neurofibromas (PN) which can cause pain and other complications[5–9]
- While a number of tools have been used to assess the impact of treatment interventions on pain severity and quality of life (QoL) in patients presenting with NF1-PN,[10] there is a scarcity of information on tools developed to evaluate the impact of treatment on QoL, specifically in adults with NF1-PN
- The plexiform neurofibromas quality of life (PlexiQoL) tool is a patient-reported outcome measure developed through rigorous qualitative methods that specifically assesses the ability of adults (aged ≥18 years) with NF1-PN to fulfill their basic needs[10–12]
  - PlexiQoL incorporates the needs-based model to provide a holistic view of the individual's current health status and ability to fulfill basic human needs, thus allowing the impacts of NF1-PN to be assessed directly rather than inferred[10,11]
- While the psychometric performance of the PlexiQoL total score was established during the development of the instrument,[10,11] data from the KOMET trial (NCT04924608; a Phase 3, multicenter, international study with a parallel, randomized, double-blind, placebo-controlled, two-arm design to assess the efficacy and safety of selumetinib in adults with NF1 and symptomatic, inoperable PN) offered an opportunity to inform the interpretability of the PlexiQoL and demonstrate that the PlexiQoL total score is fit for purpose in clinical trials of adults with NF1-PN

METHODS

- PlexiQoL comprises 18 unidimensional, dichotomous items covering five aspects of patient needs: appearance, relationships, independence, role fulfillment, and pleasure
  - Each item is given a score of 1 = true, or 0 = not true
  - A total score is obtained by summing all items, ranging between 0 and 18, with a lower score indicating better QoL
- In KOMET, PlexiQoL was assessed at baseline; Day (D) 28 of Cycles (Cs) 2, 4, 8, 12, 16, 20, and 24; and D28 of every six cycles thereafter (28-day cycles)
- PlexiQoL psychometric analysis was conducted utilizing Rasch measurement theory (RMT) and classical test theory (CTT) once all 145 patients within KOMET had the opportunity to complete Cycle (C) 12; anchor- and distribution-based methods were used to estimate the meaningful within-patient change (**Figure 1**)

Figure 1. Study analyses



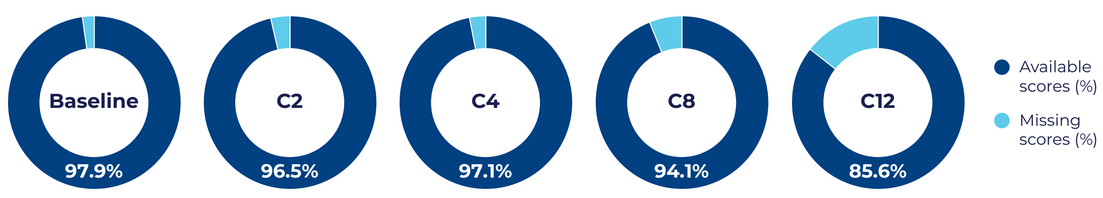
*Standard thresholds for interpretation of reliability coefficients are as follows: <0.70: unsatisfactory reliability; 0.70–0.79: modest reliability; 0.80–0.89: adequate reliability; 0.90–1.00: good reliability. Test–retest reliability estimations were conducted between baseline and C2, and between C8 and C12, in the FAS overall and in stable participants. Stable participants were defined as 1) those who reported the same responses to all EQ-5D-5L items at both cycles (first definition) and 2) those whose EQ-5D-5L VAS scores deviated by <7 in absolute value (second definition). This range was selected based on previously published research that suggested that a change in the EQ-5D-5L VAS (which is similar to the one used by the EQ-5D-5L) between 7 and 10 could be considered meaningful.[13]
C, cycle; CTT, classical test theory; EQ-5D-5L, EuroQol 5-dimension 5-level questionnaire; FAS, full analysis set; ICC, intraclass correlation coefficient; PlexiQoL, plexiform neurofibromas quality of life questionnaire; RMT, Rasch measurement theory; VAS, visual analog scale.

RESULTS

Data completeness

- Overall, 95% of assessments were complete for all PlexiQoL items between baseline and C12 (n = 647 available assessments), highlighting the good quality of the PlexiQoL data (**Figure 2**)

Figure 2. Description of PlexiQoL data completeness over the first 12 cycles (FAS, N = 145)



Available scores (n = 646): baseline (n = 142/145), C2 (n = 138/143), C4 (n = 132/136), C8 (n = 127/135), and C12 (n = 107/125). The RMT analysis is based on the item set - one assessment had two completed items that were included in the RMT analysis, but no scores that could be calculated for this assessment.
C, cycle; FAS, full analysis set; PlexiQoL, plexiform neurofibromas quality of life questionnaire; RMT, Rasch measurement theory.

- At baseline, the most endorsed items were "I feel I have no control over the plexiform(s)" (83%), "I find the plexiform(s) ugly" (74%), "I am very self-conscious about the way I look" (67%), and "My choices in life are restricted" (65%) (**Supplementary Figure S1**)
  - Mean (standard deviation [SD]) baseline PlexiQoL total score was 8.7 (4.5), and the distribution was across the full range of the scale (see **Supplementary Figure S2** for mean total scores at baseline and during follow-up)

Evaluation of the psychometric properties of the PlexiQoL using RMT analyses

- With 145 participants at baseline, RMT analysis of pooled PlexiQoL scores up to C12 (n = 647 available assessments) indicated acceptable targeting of the items for the sample, meaningful item hierarchy, acceptable fit of the items to the Rasch model, no local dependence between items, and no issues with measurement invariance across patient age and sex

Targeting of the items

- PlexiQoL items covered a large part of the continuum but not the extremes; 15% of person estimates were below the lowest item threshold and 10% were above the highest item threshold
  - This suggests that the PlexiQoL items were not able to accurately discriminate between situations where participants reported very good or very poor QoL

Meaningful item hierarchy

- A clear ranking of items was observed, illustrating a hierarchy of the needs experienced by patients with NF1-PN (**Supplementary Figure S1b**)

Item fit

- The item fit was acceptable; only a few items had statistical indicators showing deviation from the RMT (**Table 1**)

Local dependence

- Only one pair of items out of 153 items had a residual correlation >0.3: between "I cover up the plexiform(s)" and "I am very self-conscious about the way I look" (correlation = 0.35), confirming the absence of local dependence within the scale

Reliability

- The reliability of the scale estimated within the Rasch model was modest (person separation index of 0.79)

Measurement invariance

- At any given level of QoL:
  - Patients aged <30 years were more likely to endorse "I avoid looking at myself in the mirror" and "I cover up the plexiform(s)" than those aged ≥30 years, while it was the contrary for "The quality of my relationships is affected"
  - Female patients were more likely to endorse "I cover up the plexiform(s)" and "I am very self-conscious about the way I look" than male patients
- A total of 116 assessments were collected in China, 267 in Europe, 71 in Japan, and 193 in the rest of the world. Region was not found to provide participants who were more likely to endorse any item at any level of QoL

Table 1. Individual item fit of the PlexiQoL item set (FAS, N = 647 assessments)

Item	Fit residual	Chi²	p-value
I avoid crowds where possible	0.308	13.041	0.1607
I am unable to join in activities with my family and friends	1.221	20.159	0.0170
I'm losing my role in life	−3.311	29.012	0.0006
I am reluctant to leave the house	−1.890	12.297	0.1971
I avoid looking at myself in the mirror	−1.847	12.668	0.1782
I feel I have no control over the plexiform(s)	−0.276	9.349	0.4057
I find the plexiform(s) ugly	0.929	39.505	<0.0001
I take anger out on people close to me	0.606	6.570	0.6818
I avoid intimate situations	−2.466	14.281	0.1127
I find social situations stressful	−4.257	35.631	<0.0001
I can't take care of myself	0.852	4.275	0.8924
I don't like being touched	−1.686	18.464	0.0302
I feel dependent on others	1.208	12.100	0.2077
I cover up the plexiform(s)	2.709	14.273	0.1129
The quality of my relationships is affected	−3.557	20.515	0.0150
It's difficult to plan ahead	1.762	19.551	0.0209
I am very self-conscious about the way I look	0.057	28.658	0.0007
My choices in life are restricted	−1.900	18.283	0.0320

The fit residual recommended range is −2.5 to +2.5 and numbers highlighted in grey (less than −2.5) indicate items overdiscriminating with regard to the measured concept; significant values via Chi-square probabilities after Bonferroni adjustment at p<0.01 are highlighted in purple.
FAS, full analysis set; PlexiQoL, plexiform neurofibromas quality of life questionnaire.

Classical test theory

Reliability

- *Internal consistency reliability:* Cronbach's alpha score was 0.84 at baseline (N = 142) suggesting adequate (**Figure 1** legend) internal consistency reliability for group-level analyses[14]
- *Test–retest reliability:* the range in intraclass correlation coefficients (0.75–0.94) indicated modest-to-good PlexiQoL test–retest reliability (**Supplementary Table S1**)[14]

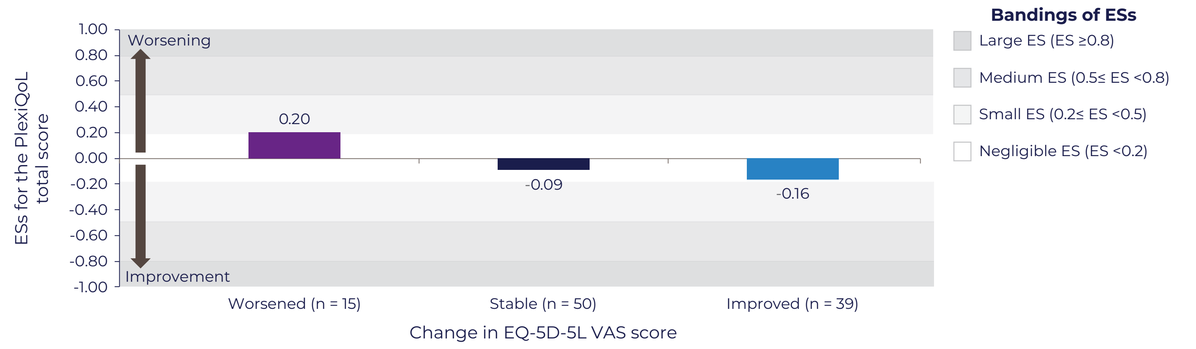
Construct validity

- PlexiQoL score showed a clear, positive association with the EQ-5D-5L items and EQ-5D-5L VAS score
  - The PlexiQoL total score was slightly greater in patients reporting higher EQ-5D-5L VAS scores (i.e., better health status overall)
  - The correlation coefficient between the two scores was −0.49, which was within the hypothesized range (≥0.4 to <0.7)
- The distribution of the PlexiQoL total score at baseline was examined according to the five EQ-5D-5L items separately
  - As expected, the PlexiQoL total score was significantly higher in patients reporting more problems walking, washing or dressing themselves, doing their usual activities, and more pain or discomfort

Ability to detect change over time

- The effect sizes of the PlexiQoL total score were assessed according to groups defined by the change in EQ-5D-5L VAS score from baseline at C12 (**Figure 3**)
  - While the direction of these effect sizes was as expected, they were all classified as negligible per Cohen's rules of thumb

Figure 3. ESs of the PlexiQoL total score according to groups defined by the change in EQ-5D-5L VAS score from baseline at C12



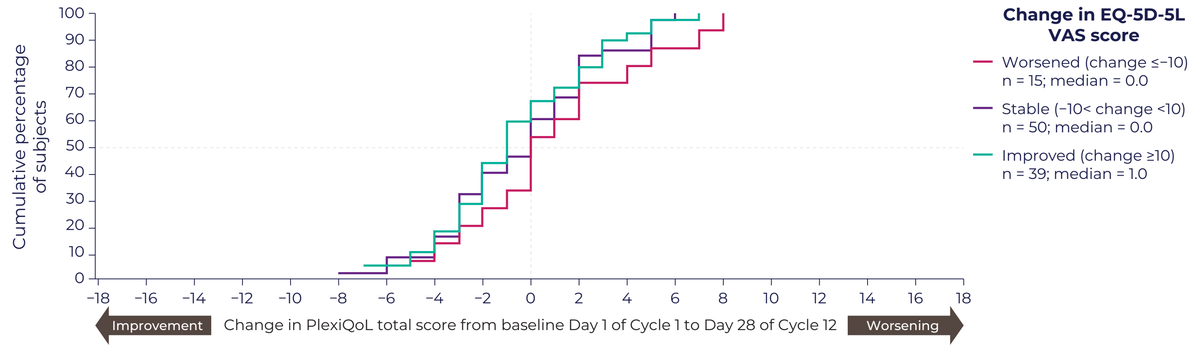
For the purpose of this analysis, participants with a change in EQ-5D-5L VAS score ≤−10 were categorized as having worsened, those with a change <10 (in absolute value) were categorized as stable, and those with a change ≥10 were categorized as having improved.
C, cycle; EQ-5D-5L, EuroQol 5-dimension 5-level questionnaire; ES, effect size; PlexiQoL, plexiform neurofibromas quality of life questionnaire; VAS, visual analog scale.

Determination of meaningful within-patient change/meaningful score difference (MSD)

Anchor-based method

- The change in PlexiQoL total score from baseline to C12 (N = 104) showed poor correlation (<0.3)[15] with both the change in EQ-5D-5L VAS score (−0.08) and the categorized change in EQ-5D-5L VAS score (−0.15) indicating the unsuitability of EQ-5D-5L VAS score as an anchor variable
  - Empirical cumulative distribution functions of change from baseline in PlexiQoL total score across EQ-5D-5L VAS score categories did not show a clear separation, reflecting the poor correlation (**Figure 4, Supplementary Table S2**)

Figure 4. Empirical cumulative distribution functions of change from baseline at C12 in PlexiQoL total score across EQ-5D-5L VAS score categories (FAS, N = 145)



For the purpose of this analysis, participants with a change in EQ-5D-5L VAS score ≤−10 were categorized as having worsened, those with a change <10 (in absolute value) were categorized as stable, and those with a change ≥10 were categorized as having improved.
EQ-5D-5L, EuroQol 5-dimension 5-level questionnaire; FAS, full analysis set; PlexiQoL, plexiform neurofibromas quality of life questionnaire; VAS, visual analog scale.

Distribution-based method

- Distribution-based estimates for meaningful change were 2.23 (0.5*SD) and 1.76 (standard error of measurement)

Triangulation

- In the absence of a suitable anchor, the meaningful score difference (MSD) estimation could only be based on the data available from the KOMET study
- Based on the distribution-based approach, a 2-point decrease is a tentative estimate of MSD for the PlexiQoL total score
- Given the scoring algorithm of the PlexiQoL, a change of 2 points in the total score would be observed when the responses to two items change (i.e., two needs are no longer met)

Acknowledgements

Medical writing support was provided by Jen Shepherd, PhD, and Suzanne Beresford, BPharm, of Helix, OPEN Health Communications, with financial support from Alexion, AstraZeneca Rare Disease, in accordance with Good Publication Practice (GPP) guidelines (www.ismpp.org/gpp-2022). This study was funded by AstraZeneca as part of an alliance between AstraZeneca and Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA (MSD).

Conflicts of interest

AA is an employee of, and holds stocks in, Alexion, AstraZeneca Rare Disease. JN is an employee of, and has stock/stock options in, Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA. JN also declares a leadership role as President of the International Society for Quality of Life Research. XY is an employee of Merck Sharp & Dohme LLC, a subsidiary of Merck & Co., Inc., Rahway, NJ, USA, and owns stock in Merck & Co., Inc., Rahway, NJ, USA. PA is employed by Alexion, AstraZeneca Rare Disease as a consultant. AR and LB are employees of Modus Outcomes. RdlRR is an employee of, and has stock/stock options in, Alexion, AstraZeneca Rare Disease. EM is an employee of AstraZeneca. AR and LB declare Alexion payment to Modus Outcomes for patient-centered outcomes research services including analyses presented in the communication.

References

1. Blakeley JO and Plotkin SR. *Neuro Oncol* 2016;18:624–638; 2. Hirbe AC and Gutmann DH. *Lancet Neurol* 2014;13:834–843; 3. Yap YS et al. *Oncotarget* 2014;5:5873–5892; 4. Wolters PL et al. *Am J Med Genet A* 2015;167A:2103–2113; 5. Miller DT et al. *Pediatrics* 2019;143:e20190660; 6. Nguyen R et al. *J Pediatr* 2011;159:652–655; 7. Tchernev G et al. *Medicine* 2016;95:e2663; 8. Gross AM et al. *N Engl J Med* 2020;382:1430–1442; 9. Yang X et al. *Pediatric Health Med Ther* 2021;11:421–428; 10. Heaney A et al. *Mol Genet Genomic Med* 2020;8:e1530; 11. Heaney A et al. *SAGE Open Med* 2019;7:2050312119829680; 12. Dhaenens BAE et al. *J Patient Rep Outcomes* 2024;8:33; 13. Pickard AS et al. *Health Qual Life Outcomes* 2007;5:70; 14. Nunally JC and Bernstein IH. *Psychometric Theory*. 3rd ed. New York, NY: McGraw-Hill Inc; 1994; 15. Hinkle DE et al. *Applied Statistics for the Behavioral Sciences.* 4th ed. Boston, MA: Houghton Mifflin Company; 2003.

Please scan this quick response (QR) code with your smartphone camera or app to obtain a copy of these materials. Alternatively, please use the link below.

https://rarediseaseresource.com/2024/ISPOR-EU-11-17/PlexiQoL_Tool_Validation_Poster

Poster | Supplementary material | Plain language summary

Poster presented at ISPOR Europe 2024; Barcelona, Spain; November 17–20, 2024.

Corresponding author: Ayo Adeyemi (Ayo.Adeyemi@alexion.com)