Claesson, M., Armitage, W. J., Byström, B., Montan, P., Samolov, B., Stenvi, U., & Lundström, M. (2017). Validation of Catquest-9SF-A Visual Disability Instrument to Evaluate Patient Function after Corneal Transplantation. *Cornea*, *36*(9), 1083-1088. https://doi.org/10.1097/ICO.0000000000001255

Validation of Catquest-9SF – a visual disability instrument to evaluate patient function after corneal transplantation

Margareta Claesson, MD PhD,[1] W John Armitage, PhD,[2,3] Berit Byström, MD PhD,[4] Per Montan, MD PhD,[5] Branka Samolov, MD PhD,[5] Ulf Stenvi, MD PhD,[1] Mats Lundström, MD PhD[6]

[1]Department of Ophthalmology, Sahlgrenska University Hospital, Mölndal, Sweden

[2]School of Clinical Sciences, University of Bristol, Bristol, UK

[3] Tissue and Eye Services, NHS Blood and Transplant, Bristol, UK

[4]Department of Clinical Sciences, Ophthalmology, Umeå University, Umeå, Sweden

[5]St Erik Eye Hospital, Stockholm, Sweden

[6]Department of Clinical Sciences, Ophthalmology, Faculty of Medicine, Lund University, Sweden

Corresponding author: Dr Margareta Claesson (margareta.l.claesson@vgregion.se)

Meeting presentation: Association for Research in Vision and Ophthalmology, Seattle, May 2016

Financial support: Funding from the Swedish Association for Local Authorities and Regions (SKL) for the Swedish Corneal Transplant Registry and institutional support. The funding organisations had no role in the design or conduct of this research.

Conflict of interest: No conflicting relationship exists for any author.

Running head: Patient reported outcome measures for corneal transplantation

Address for reprints: Department of Ophthalmology, Sahlgrenska University Hospital, Göteborgsvägen, S-431 42 Mölndal, Sweden

Abstract

Purpose. Catquest-9SF is a 9-item visual disability questionnaire developed for evaluating patient reported outcome measures (PROM) after cataract surgery. The aim of the present study was to use Rasch analysis to determine the responsiveness of Catquest-9SF for corneal transplant patients.

Methods. Corneal transplant patients where the primary reason for surgery was to improve vision were included. One group (n=199) completed the Catquest-9SF questionnaire before corneal transplantation and a second independent group (n=199) completed the questionnaire 2 years after surgery. All patients were recorded on the Swedish Cornea Registry, which provided clinical and demographic data for the study. Winsteps software v.3.91.0 (Winsteps.com, Beaverton, OR, USA) was used to assess the fit of the Catquest-9SF data to the Rasch model.

Results. Rasch analysis showed that Catquest-9SF applied to corneal transplant patients was unidimensional (infit range, 0.73 to 1.32; outfit range, 0.81 to 1.35) and therefore measured a single underlying construct (visual disability). The Rasch model explained 68.5% of raw variance. The response categories of the 9-item questionnaire were ordered and the category thresholds were well-defined. Item difficulty matched the level of patients' ability (0.36 logit difference between the means). Precision in terms of person separation (3.09) and person reliability (0.91) was good. Differential item functioning (DIF) was notable for only 1 item (satisfaction with vision), which had a DIF contrast of 1.08 logit.

Conclusions. Rasch analysis showed that Catquest-9SF is a valid instrument for measuring visual disability in corneal transplant patients where the primary reason for transplantation is to improve vision.

Since 1996, the Swedish Cornea Registry has collected clinical data from at least 90% of patients undergoing corneal transplantation in Sweden. With an addition of the two Danish centres performing keratoplasty, more than 1000 patients are added to the registry annually. The focus is on visual outcome, but data on graft survival, pre-operative risk factors and post-operative complications are also obtained. Data concerning preoperative status and the surgical intervention are collected at the time of operation. Follow up-data reporting on visual outcome, graft survival and complications, such as graft rejection, are reported at a single 2-year postoperative follow-up.[1] This time point was chosen because visual rehabilitation after penetrating keratoplasty in the majority of patients should be completed by 2 years.[2]

Clinical outcome measures, such as visual acuity, do not necessarily relate to the experience of a patient's ability to manage everyday tasks. In order to understand more about the actual benefit to patients of surgical inventions where the aim is to improve vision, several visual disability questionnaires have been developed, in particular for cataract patients. These ask patients to rate their ability to perform a range of tasks set out as a series of questions (items). Examples of visual disability questionnaires include the National Eye Institute Visual Function Questionnaire (NEI-VQF),[3, 4] Activities of Daily Vision Scale (ADVS),[5] the Visual Disability Assessment (VDA),[6] the Vision-Related Quality of Life (VR-QOL)[7] and the Visual Functioning 14 (VF-14).[8] There are also questionnaires created for single investigations.[9] These questionnaires have been shown to be sensitive to clinically meaningful changes after surgery. In many cases, however, the numeric responses to the various questions (e.g., 1=much difficulty; 2=little difficulty, etc.) have simply been added together to give a single score to describe the level of disability for the individual patient. This approach assumes that the responses given by patients to the questions depend solely on the ability of patients to perform each particular task (respondent ability) and ignores the level of difficulty of the task (item difficulty): it also implies, incorrectly, that all the questions carry equal weight in the assessment of visual disability and that the differences between the raw scores for each response category are uniformly equal. These failings can be overcome by Rasch analysis,[10, 11] which transforms raw, ordinal scores to an interval scale. Respondent ability and item difficulty are therefore sorted on the same scale, expressed in logit values. These are used to confirm the correct ordering of response categories, ranking of item difficulty and target the respondent mean ability. The logit value is the natural log-odds of a positive reaction to an item where logit = 0 is the mean item difficulty. This analysis also shows how well the questions target the patients' abilities, and determines whether the questions measure a single underlying construct or characteristic (in this instance, visual disability).

A comparison of 16 questionnaires used for cataract surgery outcomes found Catquest-9SF to be the most responsive.[12, 13] Catquest-9SF, a 9-item short form measure, was developed from the original Catquest visual disability questionnaire.[12] Catquest was introduced in 1995 to collect patient-reported outcome data for the Swedish National Cataract Registry before and at 6 months after cataract surgery.[14] Rasch analysis showed Catquest-9SF to be a valid instrument for measuring visual disability outcomes after cataract surgery.[12] We decided, therefore, to investigate whether Catquest-9SF could be applied to corneal transplant patients in order to find out the true benefit for the patients by combining clinical outcome data with patient reported outcome measures (PROM). In a recent study, a Danish translation of Catquest-9SF was used, without specific validation for corneal transplantation, for patients with Fuchs endothelial dystrophy.[15] Herein, we present a formal validation of Catquest-9SF using Rasch analysis to create a linear measure of visual disability in corneal transplant patients where the primary reason for the transplant is to improve vision.

Methods

Patients

Corneal transplant patients from 3 clinics in Sweden, namely Gothenburg, Stockholm and Umeå, were included in this study, which adhered to the tenets of the Declaration of Helsinki. Consecutive patients who were over 18 years old and capable of understanding the questionnaire were selected for this study. Because all patients had given consent for their data to be included in the Swedish Corneal Transplant Registry, the Sahlgrenska University Hospital IRB judged that further approval was not required.. Participants were included regardless of the indication for transplantation and the type of operation (penetrating or endothelial keratoplasty) provided the reason for the transplant was to improve vision. In order to maintain independence between the pre-operative and post-operative groups for the purposes of this validation study, separate groups of patients were selected to complete the questionnaire either before (Form 1) or at the routine 2-year postoperative follow up (Form 2). The total number of patients was 398, each group comprising 199 patients. The patients' characteristics are shown in Table 1. The Form 2 group included both surviving (n=164) and failed (n=35) grafts, but excluded those regrafted before the 2- year follow up was due.

Catquest-9SF

Catquest-9SF includes 7 items about perceived difficulties in performing activities in daily life and 2 global questions about, respectively, difficulties in general and satisfaction with vision. The questionnaires were answered in Swedish and Table 2 therefore shows an English translation of the items and their respective response categories. As shown in Table 2, each item has four response categories, which are allocated ordinal numeric values. The response categories included an option 'Cannot decide', which was treated as missing data. The format and item ordering in the pre- and postoperative questionnaires were identical.

Rasch analysis

The fit of the 9 items of the questionnaire (Table 2) to a single construct (visual disability) was assessed by Rasch analysis[11] using Winsteps 3.91.0 (Winsteps.com, Beaverton, OR, USA) applied with a 4-Andrich rating scale model for each question type.[16] Rasch analysis converts the raw ordinal questionnaire scores to a linear interval scale by logit transformation, thereby permitting the use of parametric statistical techniques on the questionnaire data. The scale also allows evaluation both of how well the item difficulty targets person ability and scale validity assessment, especially item and person fit to the overall construct.

Rasch analysis tests the following psychometric properties of the questionnaire:

- *Rating scale.* The Catquest-9SF items (questions) each have four response categories (see Table 2) and, therefore, three thresholds between the response probabilities. The test investigates whether the category thresholds are ordered and well-defined such that patients are able to distinguish adequately between neighbouring response categories; for example, being able to distinguish in Item 2 between “Fairly satisfied” and “Very satisfied”.
- *Person separation and reliability.* These show the precision of the instrument, which means how well the instrument is able to discriminate different levels of person ability. A low person separation indicates that the instrument may not be sensitive enough to distinguish, in this case, between high and low visual disabilities. Person reliability shows how reproducible the responses would be if the same individuals were to be given another set of questions dealing with the same construct.
- *Item fit statistics.* Infit and outfit mean squares (MNSQ) indicate variation between the observed response patterns and those predicted by the Rasch model. They should have a value of 1 (acceptable range 0.7-1.3). Infit mean square >1 indicates more variation in the observed data than predicted by the model. Outfit mean square <1 indicates less variation in the observed response pattern than predicted.

- *Targeting.* How well the distribution of items matches the range of person abilities such that there are meaningful items for both the more able and less able persons. The means of the items and the persons should be similar. However, SDs are seldom similar because the spread of respondent ability is usually much greater than the spread of item difficulty.
- *Unidimensionality.* The measurement of a single underlying characteristic (i.e., visual disability), which is essential for the creation of a summary score.
- *Differential item functioning (DIF).* An indication of bias between groups of patients owing to, for example, age, gender or co-morbidity. For this study, the DIF evaluated patient responses before and after transplantation.

Results

Rasch analysis

The characteristics of the pre- and post-operative groups were similar for age, gender ratio and distribution of indications (Table 1); but, as anticipated, there was a marked improvement in visual acuity two years after transplantation (p<0.0001). Pre-operatively, only 4% of patients had a visual acuity ≥0.5 (20/40) compared with 43% post-operatively, and the respective percentages of patients with poor visual acuity (≤0.2 [20/100]) were 80% before and 38% after transplantation.

The category probability curves in Figure 1 show that the response categories were ordered. The category thresholds were well-defined (distance between thresholds ≥1.4 and <5.0 logits) and increased monotonically across the rating scale. Table 2 shows the item fit characteristics. The instrument distinguished well between different levels of patients' ability, the real person separation was 3.09 (expected ≥2.5) and person separation reliability was 0.91 (expected ≥0.85). The items in the instrument fitted the Rasch model expectation, (infit and outfit MNSQ should be between 0.7 and 1.3): infit and outfit MNSQ were ≥0.73 for all items and only one item had MNSQ >1.3 (Item 6:

infit MNSQ 1.32, outfit MNSQ 1.35). Principal components analysis of the residuals showed that 68.5% of the variance was explained by the measures. The unexplained variance explained by the first contrast was 1.7 Eigenvalue units (5.8%). This demonstrated that the instrument was unidimensional and measured only a single characteristic (i.e., visual disability). The responses of the pre- and post-operative subgroups showed a noticeable difference only for Item 2 (satisfaction with vision), which had a DIF contrast of 1.08 logits (expected <0.5 logits). Item difficulty matched the level of participants' ability.

The Item-Person map (Figure 2) shows that the instrument was well-targeted since the distribution of items matches the range of patients' abilities. The means of the distributions differed by only 0.36 logits (expected ≤1.0 logits) and the patients' locations were evenly distributed.

Figure 3 shows a preliminary example of the use of Catquest-9SF where each patient answered the questionnaire both before surgery and at the 2-year follow up. The figure compares pre- and post-operative Rasch scores for the indications keratoconus, Fuchs endothelial dystrophy, pseudophakic bullous keratopathy and a mixed group of other indications. For all the indications the Rasch score is lower after the corneal transplantation compared with before surgery, showing improved visual ability.

Discussion

The clinical outcome after corneal transplantation, be it penetrating or lamellar, depends mainly on recipient factors, such as indication and pre-operative risk factors; however, actual benefit as perceived by patients is not necessarily revealed by clinical outcome measures and may vary between individuals with seemingly similar clinical outcomes. Many questionnaires have been developed to collect patient reported outcome measures (PROM) mainly from patients before and after cataract surgery.[13] In our search for a suitable questionnaire for patients undergoing corneal transplantation we chose to validate Catquest-9SF. It is well established for cataract surgery, short (only 9 questions) and therefore well suited for clinical practice. It has also been shown to be highly responsive to cataract surgery.[12] The application of Rasch analysis provides a greater insight into internal consistency through the fit of the items to the model and the scoring of patient ability on a valid interval scale. This improves the precision of the instrument and results in more meaningful interpretation of scoring and reduces the sample size required to find significant differences when applied to actual outcome studies.

It was important in this study to have independent groups of patients answering the pre- and post-operative questionnaires to avoid introducing bias that would have undermined the validation exercise. Clearly, in a study to investigate the benefit of surgery, each patient would need to complete the form before and at a specific time point after the transplant operation (e.g., see Figure 3).

In our study, Rasch analysis showed that the responses to each question were well defined (i.e., not overlapping). The instrument was unidimensional: it measured a single underlying construct (visual disability) since nearly 70% of the variability in responses could be explained by the Rasch model and the Eigenvalue of the first contrast was <2.0. The unexplained 30% of variance includes randomness in the Rasch model and indicates a low level of multidimensionality in the instrument. The analysis showed a good measurement precision with high sensitivity and reproducibility. The items in the instrument fit the Rasch model expectations, which means that the difficulty level of the questions was appropriate for the patients. The responses of the pre- and post-operative subgroups showed a noticeable difference (DIF) only for Item 2. This item concerns patient satisfaction with vision and, since the percentage of patients with good visual acuity was rather higher in the post-operative group than in the pre-operative group (P<0.0001), a difference in patient response to this question between the two groups was not unexpected. The improvement in visual acuity depends on the indication for transplantation;[1] however, the distribution of indications was similar in the two groups (p=0.5). Finally, item difficulty matched the level of patients' ability, which is shown as an even spread of results in the Item-Person map (Figure 2).

Rasch analysis showed that Catquest-9SF is a very suitable questionnaire for patient reported outcome measurements in patients undergoing corneal transplantation where the reason for the operation is to improve vision. In a preliminary analysis where each patient answered the questionnaire both before and two years after transplantation, the Rasch score was lower after the transplantation, meaning improved visual ability, for all the indications (Figure 3).[17] For comparison, pre- and post-operative Rasch scores for cataract patients have been reported as 0.28 logits (range -5.79 to 5.38) and -3.61 (range -5.79 to 0.78), respectively.[18] Visser et al.[19] give further practical guidance on using Catquest-9SF, including a quick access table for the clinical interpretation of Catquest-9SF scores.

In future studies of corneal transplant patients, additional factors such as gender, age and co-morbidity in the grafted or fellow eye should be included. This instrument could be of particular use in controlled studies comparing different keratoplasty techniques. Based on our validation study, the Swedish Cornea Registry steering group has endorsed the use of Catquest-9SF to complement clinical outcome measures with patient reported outcome measures and broaden the overall understanding of the factors influencing corneal graft outcome.

References

1. Claesson M, Armitage WJ, Fagerholm P et al. Visual outcome in corneal grafts: a preliminary analysis of the Swedish Corneal Transplant Register. *Brit J Ophthalmol* 2002;86:174-180.
2. Claesson M, Armitage WJ. Ten-year follow-up of graft survival and visual outcome after penetrating keratoplasty in Sweden. *Cornea* 2009;28:1124-1129.
3. Mangione CM, Lee PP, Pitts J et al. Psychometric properties of the National Eye Institute Visual Function Questionnaire (NEI-VFQ). NEI-VFQ Field Test Investigators. *Arch Ophthalmol* 1998;116:1496-1504.
4. Pesudovs K, Gothwal VK, Wright T et al.. Remediating serious flaws in the National Eye Institute Visual Function Questionnaire. *Journal of Cataract & Refractive Surgery* 2010;36:718-732.
5. Mangione CM, Phillips RS, Lawrence MG et al.. Improved visual function and attenuation of declines in health-related quality of life after cataract extraction. *Arch Ophthalmol* 1994;112:1419-1425.
6. Pesudovs K, Coster DJ. An instrument for assessment of subjective visual disability in cataract patients. *Brit J Ophthalmol* 1998;82:617-624.
7. Frost A , Eachus J, Sparrow J, et al. Vision-related quality of life impairment in an elderly UK population: associations with age, sex, social class and material deprivation. *Eye* 2001;15:739-744.
8. Steinberg EP, Tielsch JM, Schein OD, et al. The VF-14. An index of functional impairment in patients with cataract. *Arch Ophthalmol* 1994;112:630-638.
9. Uiters E, van den Borne B, van der Horst FG et al.. Patient satisfaction after corneal transplantation. *Cornea* 2001;20:687-694.
10. Rasch G. *Probabilistic models for some intelligence attainment tests*. Copenhagen: Danmarks Pædagogiske Institut; 1960.
11. Bond TG, Fox CM. *Applying the Rasch model, 3rd edn.* New York: Routledge; 2015.

12. Lundstrom M, Pesudovs K. Catquest-9SF patient outcomes questionnaire: nine-item short-form Rasch-scaled revision of the Catquest questionnaire. *Journal of Cataract & Refractive Surgery* 2009;35:504-513.

13. McAlinden C, Gothwal VK, Khadka J et al.. A head-to-head comparison of 16 cataract surgery outcome questionnaires. *Ophthalmology* 2011;118:2374-2381.

14. Lundstrom M, Stenevi U, Thorburn W. The Swedish National Cataract Register: A 9-year review. *Acta Ophthalmol Scand* 2002;80:248-257.

15. Nielsen E, Ivarsen A, Kristensen S et al.. Fuchs' endothelial corneal dystrophy: a controlled prospective study on visual recovery after endothelial keratoplasty. *Acta ophthalmologica* 2016;94:780-787.

16. Andrich D. Application of a psychometric rating model to ordered categories which are scored with successive integers. *Applied Psychological Measurement* 1978;2:581-594.

17. Claesson M, Armitage WJ, Lundström M. Validation of Catquest-9SF – A visual disbility instrument for corneal transplantation. *Invest Ophthalmol Vis Sci* 2016;57:ARVO E-Abstract 1225

18. Lundström M, Llovet F, Llovet A et al. Validation of the Spanish Catquest-9SF in patients achieving a monofocal or trifocal intraocular lens. *J Cataract Refract Surg* 2016;42:1791-1796.19. Visser MS, Dieleman M, Klijn S et al. Validation, norm scores and test-retest for the Dutch Catquest-9SF. *Acta Ophthalmol*

Figure legends

Figure 1. Category probability curves: A, items 1 and 3-9; B, item 2. Each line represents a response category: for A, red is 'yes, very great difficulty, blue is 'yes, great difficulty'' lilac is 'yes, some difficulty and black is 'no, no difficulty'; for B, red is 'very dissatisfied', blue is 'rather dissatisfied', lilac is 'fairly satisfied' and black is 'very satisfied'. The intersections (category thresholds) where there is an equal probability of a patient choosing one of two adjacent response categories, were evenly spaced, differing between 1.4 and 5.0 logits.

Figure 2. Item-Person map showing the extent to which item difficulty matches the level of patients' ability. Key: each '#' represents 3 persons and each '.' 1-2 persons; M, S and T represent, respectively, the mean, one standard deviation and two standard deviations; each number on the right-hand side of the vertical line identifies the relevant item. The instrument is well-targeted since the item means and person means were separated by only 0.36 logits.

Figure 3. Preliminary results from the application of Catquest-9SF to a single group of patients before and after corneal transplantation for different indications (KCN, keratoconus; FED, Fuchs endothelial dystrophy; PBK, pseudophakic bullous keratopathy). The boxplots show the median Rasch scores, interquartile range, minimum and maximum scores, and outliers (blue, preoperative scores; green, postoperative scores). The lower the score the better a patient's perceived visual ability.

Table 1. Patient characteristics. Pre-operative group (Form 1, n-199) compared with post-operative group (Form 2, n=199) 2 years after corneal transplantation.

	Form 1	Form 2	p
Age (years, mean [SD])	68 (15)	68 (15)	0.7
Gender (M:F)	51:49	42:58	0.1
Indication[1]			
KCN	28 (14%)	20 (10%)	
FED	78 (39%)	83 (42%)	
PBK	39 (20%)	32 (16%)	
Regraft	35 (18%)	38 (19%)	
Other	19 (10%)	26 (13%)	0.5
Visual acuity[2]	Pre-operative	Post-operative	
Good (≥0.5)	7 (4%)	86 (43%)	
Moderate (0.3-0.4)	32 (16%)	37 (19%)	
Poor (≤0.2)	160 (80%)	76 (38%)	<0.0001

Notes:

[1]KCN, keratoconus; FED, Fuchs endothelial dystrophy; PBK, pseudophakic bullous keratopathy

[2]Snellen visual acuity in decimal notation: 0.5 = 20/40; 0.4 = 20/50; 0.3 = 20/67; 0.2 = 20/100

Table 2. Catquest-9SF questionnaire (translated from Swedish) and item fit characteristics

Item	Item calibration (standard error)	Infit MNSQ[2]	Outfit MNSQ[2]	DIF[3] (pre-op to post-op)
Global assessment items[1]				
1. Do you find that your sight at present in some way causes you difficulty in your daily life?	-0.37 (0.09)	0.73	0.81	0.12
2. Are you satisfied or dissatisfied with your sight at present?	-1.27 (0.10)	0.90	0.95	-1.08
Difficulty items[1]				
Do you have difficulty with the following activities because of your sight?				
3. Reading text in newspapers	-0.13 (0.09)	0.92	0.90	0
4. Recognizing faces of people you meet	0.98 (0.01)	1.28	1.20	-0.19
5. Seeing prices of goods when shopping	-0.06 (0.09)	0.86	0.88	0.39
6. Seeing to walk on uneven surfaces	0.49 (0.10)	1.32	1.35	0.61
7. Seeing to do handicrafts	-0.48 (0.10)	0.85	0.79	0.30
8. Reading subtitles on TV	0.51 (0.09)	1.01	1.02	0.07
9. Seeing to engage in an activity/hobby that you are interested in	0.32 (0.10)	1.06	0.96	-0.14

Notes:

[1]Response categories:

Items 1 and 3-9

Yes, very great difficulty; Yes, great difficulty; Yes, some difficulty; No, no difficulty.

Item 2

Very dissatisfied; Rather dissatisfied; Fairly satisfied; Very satisfied.

Patients can also select a ‘cannot decide’ option for any of the items, which is treated as missing data.

[2]MNSQ, weighted mean square statistics

[3]DIF, Differential Item Functioning

Figure 1

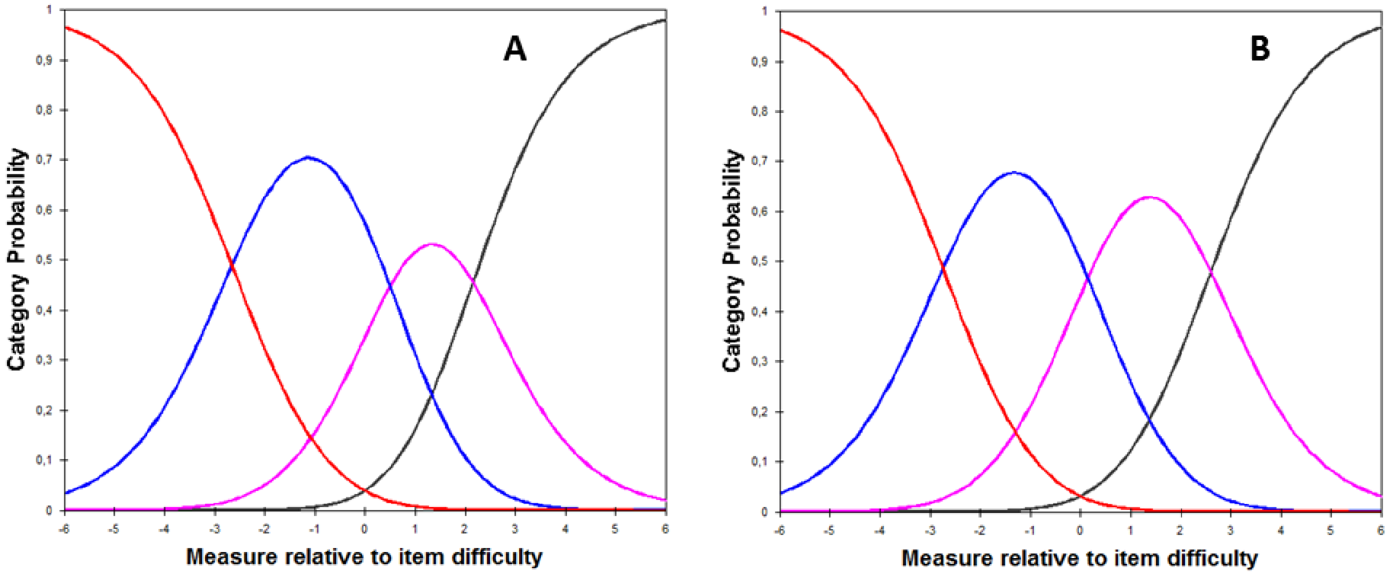

A
B
Category Probability
Measure relative to item difficulty
1
0,9
0,8
0,7
0,6
0,5
0,4
0,3
0,2
0,1
0
-6
-5
-4
-3
-2
-1
0
1
2
3
4
5
6

Figure 2

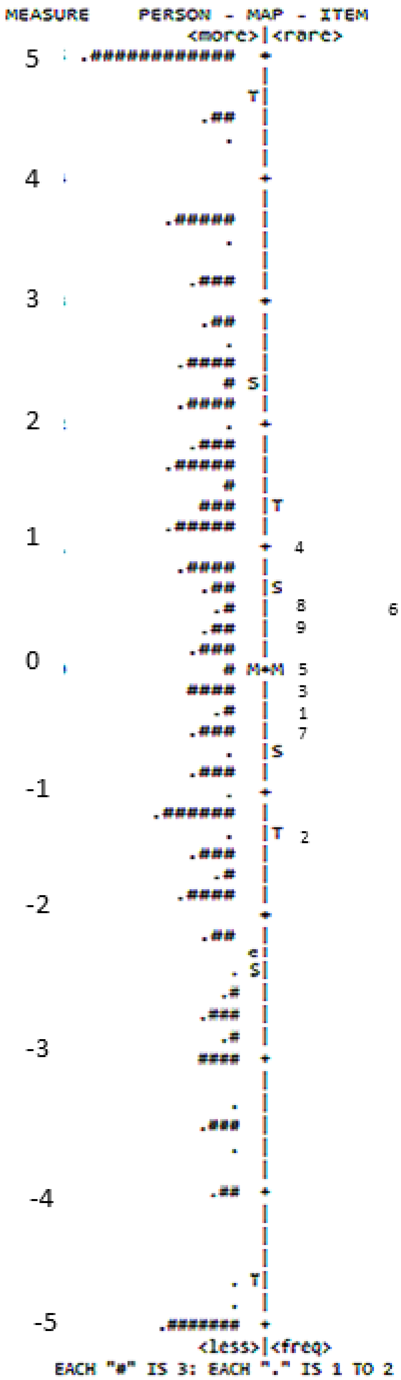
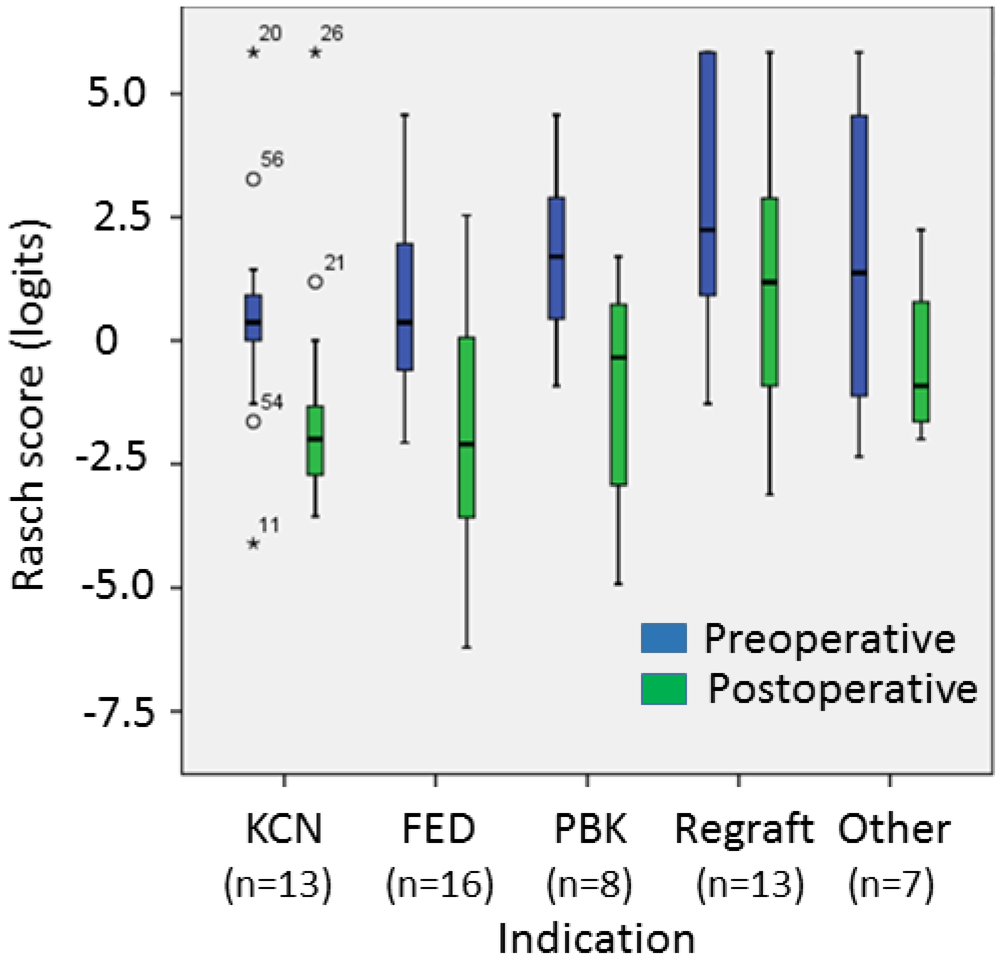

MEASURE PERSON - MAP - ITEM
<more>|<rare>
5
4
3
2
1
0
-1
-2
-3
-4
-5
<less>|<freq>
EACH "#" IS 3: EACH "." IS 1 TO 2

Figure 3



Rasch score (logits)
5.0
2.5
0
-2.5
-5.0
-7.5
20
26
56
21
54
11
Preoperative
Postoperative
KCN
(n=13)
FED
(n=16)
PBK
(n=8)
Regraft
(n=13)
Other
(n=7)
Indication