



ORIGINAL ARTICLE

The facioscapulohumeral muscular dystrophy Rasch-built overall disability scale (FSHD-RODS)

Karlien Mul¹  | Tatiana Hamadeh²  | Corinne G. C. Horlings¹ | Rabi Tawil³ | Jeffrey M. Statland⁴ | Sabrina Sacconi⁵ | Alastair J. Corbett⁶ | Nicol C. Voermans¹ | Catharina G. Faber⁷ | Baziel G. M. van Engelen¹ | Ingemar S. J. Merkies^{2,7}

¹Department of Neurology, Donders Institute for Brain, Cognition and Behaviour, Radboud University Medical Center, Nijmegen, The Netherlands

²Department of Neurology, Curaçao Medical Center, Willemstad, Curaçao

³Department of Neurology, University of Rochester Medical Center, Rochester, NY, USA

⁴Department of Neurology, University of Kansas Medical Center, Kansas City, KS, USA

⁵Centre de référence des Maladies Neuromusculaires, Nice, France

⁶Department of Neurology, Concord Hospital Medical Center, Concord, NSW, Australia

⁷Department of Neurology, Maastricht University Medical Center, Maastricht, The Netherlands

Correspondence

Karlien Mul, Department of Neurology (910), Radboud University Medical Center, PO Box 9101, 6500 HB Nijmegen, The Netherlands.
Email: karlien.mul@radboudumc.nl

Funding information

Prinses Beatrix Spierfonds, Grant/Award Number: W.OR12-22

Abstract

Background and objectives: Facioscapulohumeral muscular dystrophy (FHD) is a debilitating inherited muscle disease for which various therapeutic strategies are being investigated. Thus far, little attention has been given in FHD to the development of scientifically sound outcome measures fulfilling regulatory authority requirements. The aim of this study was to design a patient-reported Rasch-built interval scale on activity and participation for FHD.

Methods: A pre-phase FHD-Rasch-built overall disability scale (pre-FHD-RODS; consisting of 159 activity/participation items), based on the World Health Organization international classification of disease-related functional consequences was completed by 762 FHD patients (Netherlands: $n = 171$; UK: $n = 287$; United States: $n = 221$; France: $n = 52$; Australia: $n = 32$). A proportion of the patient cohort completed it twice ($n = 230$; interval 2–4 weeks; reliability studies). The pre-FHD-RODS was subjected to Rasch analyses to create a model fulfilling its requirements. Validity studies were performed through correlation with the motor function measure.

Results: The pre-FHD-RODS did not meet the Rasch model expectations. Based on determinants such as misfit statistics and misfit residuals, differential item functioning, and local dependency, we systematically removed items until a final 38-item (originating from 32 items; six items split) FHD-RODS was constructed achieving Rasch model expectations. Adequate test-retest reliability and (cross-cultural and external) validity scores were obtained.

Conclusions: The FHD-RODS is a disease-specific interval measure suitable for detecting activity and participation restrictions in patients with FHD with good item/person reliability and validity scores. The use of this scale is recommended in the near future, to determine the functional deterioration slope in FHD per year as a preparation for the upcoming clinical intervention trials in FHD.

KEYWORDS

activity and participation, facioscapulohumeral dystrophy, FHD, outcome research, Rasch-built disability scale, reliability, validity

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *European Journal of Neurology* published by John Wiley & Sons Ltd on behalf of European Academy of Neurology

INTRODUCTION

Facioscapulohumeral muscular dystrophy (FSHD) is an inherited muscle disorder that is characterized by slowly progressive muscle weakness and wasting of facial and shoulder girdle muscles and, in later stages, the trunk and leg muscles [1]. Currently there is no cure or pharmacological treatment available for FSHD, but since the discovery of its (epi)genetic mechanism, various therapeutic strategies are being investigated [2]. Multiple pharmaceutical companies have active drug development programs for FSHD, clinical trials are currently ongoing and more are expected to be initiated within the next few years [3,4].

Especially for late-phase clinical trials on FSHD, there is now an urgent need for clinical outcome measures that indicate how a patient 'feels, functions or survives' [5]. The need for patient-relevant outcome measures has been emphasized both by the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA) and by the international FSHD research community [3,6–8]. Thus far, limited attention has been given in FSHD to the development of scientifically sound outcome measures fulfilling regulatory authority requirements, with the focus on patient-reported outcomes. To date only one patient-reported disease-specific ordinal scale is available that measures the FSHD disease burden (the FSHD-health index) and a disease-specific physician-reported functional outcome measure is under further development (the FSHD-composite outcome measure) [3,9–11]. More generic outcome measures, even the ones designed to measure patients with neuromuscular disorders, are often not optimally suited to measure FSHD patients. For example, for the motor function measure (MFM), an examiner-reported scale that assesses severity of motor deficits, was shown to have a large ceiling effect in FSHD and showed a lack of items to adequately capture various degrees of motor deficits across the entire severity spectrum [12–14].

Additionally, nearly all available outcome measures are ordinal scales which provide a systemic ordering instead of a numerical value. Consequently, these scales are not suited for parametric statistical testing and pose a risk of misinterpreting clinical trial results [15,16]. To overcome the shortcomings of ordinal scales, techniques such as Rasch modeling can be used to create interval scales [17].

To fill the gap of clinical outcome measures for (late-phase) clinical trials in FSHD [10], the primary aim of this paper was to present the development of a Rasch-built overall disability scale (RODS) specifically designed for patients with FSHD and to examine its scientific soundness [18,19].

METHODS

Patients, eligibility and ethical approval

A total of 762 patients with FSHD were recruited between 2015 and 2018. Dutch patients were recruited at the Department of Neurology of the Radboud University Medical Center, Nijmegen,

the Netherlands, as part of a large observational cohort study (FSHD-FOCUS study) [20,21]. UK, US and French patients were recruited through the respective national FSHD registries [9,22–24]. Australian patients were recruited through the neuromuscular clinic of the Concord Hospital Medical Center, Concord, Australia. Eligibility was based on the following criteria: age 18 years and older, genetically confirmed FSHD and written informed consent before study enrollment. Basic characteristics such as age, sex, and country of origin were collected. This study was conducted according to the principles of the Declaration of Helsinki (version October 2013) and in accordance with the Medical Research Involving Human Subjects Act (WMO). The local medical ethics committee in the Netherlands (CMO region Arnhem-Nijmegen) approved the study protocol.

Questionnaire development

As previously reported, published standardized requirements for scale development were applied to create the FSHD disease-specific activity and participation scale [18,19,25].

The pre-phase FSHD-RODS questionnaire was composed of a list of 146 previously selected items from the World Health Organization International Classification of Functioning, Disability and Health (WHO-ICF) item list including items both on activities and participation [25–27]. Another 13 items were added by the FSHD experts from the Dutch study team, which were expected to be relevant specifically for FSHD patients, for example, items relating to facial functioning. This resulted in a pre-phase total of 159 activity and participation items, scoring each item as "0" unable to perform, "1" able to perform, but with difficulty, or "2" easily performed, without difficulty [25]. An item was scored "3" if it was not applicable to the patient and coded as "missing" for the scale development.

Additional outcome measure

The Dutch patients completed the MFM [28]. This is a 32-item instrument that assesses the severity of the motor deficit in neuromuscular disorders and was validated in a cohort of patients with a variety of neuromuscular disorders including FSHD. Items are scored from 0 to 3, with higher scores reflecting better clinical health. For this study, we used the MFM sum score of ordinal values for the external construct convergent validity of the final FSHD-RODS.

Assessment procedure

Standardized written instructions were given to patients to complete the pre-FSHD-RODS. In case clarification was required, patients could contact the researchers by email. In brief: patients were instructed that the answers given to each item should solely

be related to the possible impact of their FSHD, rather than another concomitant disease (e.g., arthritis). Patients were also instructed that there are no right or wrong answers, but it is how they feel that is important. In case of any doubt completing a task, the patient is requested to choose an answer that reflects the best of his/her personal performance ability related to the FSHD. In case an item is not relevant to the situation (e.g., "I never dance"), the patient is requested to choose an answer to the best of his/her personal judgment of being able or not to perform such an item. "Not applicable" should be used only in exceptional cases when the patient has a real problem choosing any of the other options. If the patient uses a device (e.g., adapted cutlery or a walking device) to complete an item, the patient is requested to score "possible to perform, but with difficulty" when being able to execute. In case the patient was not able to complete a task, even with the help of a device, the patient is instructed to score "impossible to perform".

Dutch and UK patients were requested to complete the pre-FSHD-RODS questionnaire twice (test-retest study; interval 2–4 weeks). The MFM (for external validity of the final FSHD-RODS) was completed once by the Dutch patients.

Rasch analyses

The pre-phase FSHD-RODS was subjected to the Rasch Unidimensional Measurement Model (RUMM2030 software) to determine whether model expectations would be met [29,30]. Numerous educational papers on Rasch have been described [17,31], including one for the neurology community [25]. In brief, the Rasch technique enables the transition of obtained ordinal scores to interval scores. Analyses were performed to obtain a final constructed FSHD specific scale (FSHD-RODS) that would fulfill Rasch model expectations, such as proper fit statistical properties, no differential item functioning, no disordered thresholds, unidimensionality, and no local dependency [17,31,32]. The Partial Credit Model was set as the default. For the purposes of the present study, the following person factors were introduced as part of the scale's construction: age categories (<50 years, 50 to 65 years, >65 years (~equally distributed), sex (female vs. male), and country (the Netherlands vs. the United Kingdom vs. the United States vs. France vs. Australia).

To test for unidimensionality the independent *t*-test approach was used [33]. In this approach, two sets of items are determined from a principal component analysis of residuals.

If the scale is unidimensional, any subset of items within the scale should provide the same estimate of person ability. Therefore, person estimates are derived from and compared between the two most divergent sets of items using a series of *t*-tests. If the number of significant *t*-tests at the 0.05-level is below 5%, or the lower bound of a binomial 95% confidence interval (CI) of the observed proportion overlaps 5%, the scale would show unidimensionality.

Reliability and validity studies

Internal reliability of the final FSHD-RODS was examined by determining the person separation index (PSI), which should preferably be >0.9 for clinical proper discriminatory ability [32].

Test-retest reliability studies (items' hierarchy locations and patients' ability locations [Logits]) were also performed through graph analyses determining the consistency of the final scale created using 95% CIs [34]. Reliability was estimated through linear regression studies (expressed as R^2). External validity of the final FSHD-RODS was determined through correlation with the MFM sum scores [14]. Further analyses were undertaken using Stata 13.0 for Windows. The *p* value was adjusted throughout the analyses and, if needed, Bonferroni multiple testing corrections were applied [35].

RESULTS

Study population and data quality control

Initially, 762 records on 159 items were available. Through data quality control, using an arbitrarily taken cut-off of >10% missing scores on the pre-FSHD-RODS records, we omitted 13 items and 48 persons' records. In addition, two items were omitted based on low face validity, leaving 144 items and 714 patient records for further investigations. The mean (standard deviation [SD]; range) age was 54.4 (15.1; 18–87) years and there were 372 women (49%) in the cohort. A total of 153 of the patients (21.4%) were from the Netherlands, 268 (37.5%) were from the United Kingdom, 45 (6.3%) were from France, 32 were from Australia (4.5%), and 216 were from the United States (30.3%). For the external validity studies, only the Dutch data were available. For the reliability studies, a total of 230 records were available.

Initial Rasch analyses on the pre-phase FSHD-RODS

The pre-phase 144-item FSHD-RODS did not meet Rasch model requirements. The item fit (mean fit residuals -0.535 , $SD = 2.470$) and person fit (mean fit residuals -0.143 , $SD = 1.681$) residual statistics deviated substantially from model expectations, as was the case for the item–trait interaction, which showed a significant chi-square probability ($p < 0.00001$), thus, with no invariance. In addition, a proportion of 0.13 of the *t*-tests fell outside the ± 1.96 range, indicating multidimensionality. No item showed disordered thresholds.

Data handling of the pre-phase FSHD-RODS to fit Rasch modeling

Step 1

Throughout the following steps, the class intervals were continuously checked for their magnitude appropriateness. A total of 57 items

demonstrated misfit statistics and/or fit residuals exceeding ± 2.5 and were removed stepwise (87 items remaining). No patient records were removed.

Step 2

A total of 13 items showed cross-cultural misfit (uniform differential item functioning on country) and eight items showed uniform differential item functioning on sex. These 21 items were systematically evaluated: that is, each item was checked regarding its location on the ruler as well as its corresponding characteristic curve (item characteristic curve [ICC]) in relation to the projected class intervals (evaluating possible proper, under- or overdiscrimination) to determine whether the item could be removed without affecting the item and corresponding threshold distribution continuum of the model under construct. These 21 items were removed stepwise (66 items remaining; some examples of item bias are provided in Figure S1).

Step 3

Various significant correlations were seen between item residuals. All pairs of items with a correlation ≥ 0.30 were evaluated starting with the highest correlations (>0.7 , >0.6 , up to >0.30). Of each item set, the item demonstrating the least relevance in relation to the generally known clinical FSHD picture or the most over- or underdiscrimination on its ICC was removed. Eventually, a total of 22 items were removed one by one (44 items remaining). Additionally, four items showing less clinical face validity (e.g., the ability to shake out table cloth) were removed (40 items remaining).

Step 4

Statistics were improving, but needed further adjustments. We subsequently lowered the p value for the fit residuals to 0.01, enabling

eight more items to be additionally removed (32 items remaining). Summary statistics improved (items' mean fit residuals: -0.362 , SD = 1.153 and persons' mean fit residuals: -0.241 , SD = 0.841, chi-squared item-trait interaction: $p = 0.05$).

Step 5

Five items (ability to run, jump, dance, walk outdoors >1 km, and ability to walk in the dunes) showed uniform differential item functioning on age. Particularly patients younger than 50 years deviated significantly from the other two age categories (Figure 1). Taking the items' and corresponding thresholds' distribution into account, we concluded that the location of these five items was important enough for a more adequate item distribution pattern. Hence, these items were not omitted, but splitting these items was taken into account. Since RUMM2030 software does not enable examination for unidimensionality after having split an item, we examined this entity before splitting any item. Two subsets of items were formed (four most positively loaded vs. four most negatively loaded) and an acceptable trend towards unidimensionality was obtained through first principal components analyses (independent t -tests between the two groups of items: proportion of significant t -tests 0.06 [95% CI 0.04–0.08]; see also Table 1). Subsequently, these abovementioned five items were split based on item bias on age categories (for all items: patients aged <50 years experienced these items as being easier to execute compared to patients aged ≥ 50 years). Additionally, item 31 (ability to dance) for patients aged up to 50 years was split based on sex into item₃₁ <50 years, males versus item₃₁ <50 years, females (32 items still remaining, but now with 38 corresponding inquiries). Finally, we succeeded in constructing the FSHD-specific Rasch-built overall disability scale (FSHD-RODS; 38 inquiries originating from 32 remaining items) that acceptably met Rasch model expectations (item fit residuals: mean -0.389 , SD = 0.940; person fit residuals: mean -0.231 , SD = 0.788; item-trait chi-square: p value = 0.31, degrees of freedom: 342; Figure 2), except for three significant residual

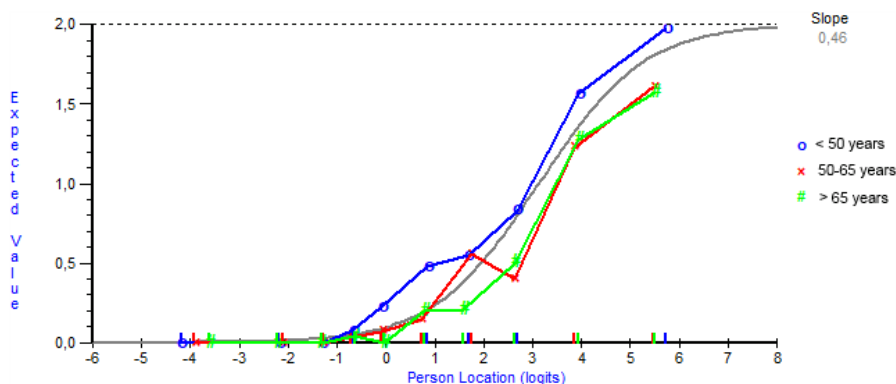


FIGURE 1 Item characteristic curve (ICC; grey line) for item 19 (ability to jump) as an example of item bias on factor age. ICC for item 19 (jump) showing uniform differential item functioning on age. The graph shows how patients aged <50 years (blue line) find this item easier to execute compared to the other two age categories (red [50 to 65 years] and green [>65 years] lines). Based on these findings, this item was split into item₁₉ <50 years versus item₁₉ ≥ 50 years [Colour figure can be viewed at wileyonlinelibrary.com]

TABLE 1 Rasch analysis findings during the construction of the activity and participation scale for patients with facioscapulohumeral muscular dystrophy

Analysis ↓ number	Item Fit Residuals		Person Fit Residuals		Item-trait Chi ² -probability			Unidimensionality t-tests (95% CI) ^b
	Mean	SD	Mean	SD	DF	p	PSI	
1	-0.535	2.470	-0.143	1.681	1296	<0.000001	0.99	0.13 (0.12–0.15)
20	-0.290	1.346	-0.312	1.355	1008	<0.000001	0.99	0.31 (0.30–0.33)
30	-0.156	1.127	-0.305	1.244	828	<0.000001	0.99	0.14 (0.12–0.15)
40	-0.182	1.028	-0.298	1.131	621	<0.000001	0.98	0.14 (0.12–0.15)
50	-0.337	1.210	-0.235	0.921	396	0.0001	0.98	0.09 (0.07–0.10)
59	-0.362	1.153	-0.241	0.841	306	0.05	0.97	0.06 (0.04–0.08)
62 (Final)	-0.389	0.940	-0.231	0.788	342	0.31	0.98	Unable to test ^a

Abbreviations: CI, confidence interval; DF, degrees of freedom; PSI, person separation index; SD, standard deviation.

^aUnable to test for unidimensionality in RUMM2030 after splitting an item.

^bProportion of significant t-tests.

correlations being still present. However, for maintaining a more appropriate item and corresponding threshold distribution, we decided to keep these items in the model. The item “ability to fill in a form” turned out to be the easiest item to execute, while “ability to run and being 50 years or older” was the most difficult item to accomplish (Figure 2). The item difficulty ranged from -4.023 to 4.304 Logits. Patient location ranged from -7.668 to 6.962. Two of the 714 patients (0.3%) showed a floor effect and 47 (6.6%) showed a ceiling effect on the final model.

Validity and reliability studies

The final 32-item (38 inquiries) FSHD-RODS demonstrated good discriminative validity when correlated with the MFM (Figure 3; Spearman's correlation coefficient 0.86). A seemingly smaller ceiling effect was seen with the FSHD-RODS when compared to the MFM sum scores. Internal reliability for the final FSHD-RODS was robust (PSI: 0.97). Also, excellent test-retest reliability scores were obtained for item hierarchy and patient abilities (Figure 4). The final FSHD-RODS is presented in Table 2.

DISCUSSION

This paper presents a patient-reported Rasch-built activity and participation interval scale specifically designed for patients with facioscapulohumeral muscular dystrophy, the FSHD-RODS. The final FSHD-RODS demonstrated adequate discriminatory validity and reliability scores including cross-cultural validity. It is easy to use and can be completed in a matter of only minutes.

Most of the currently used clinical outcome measures in research on FSHD have serious limitations. Nearly all outcome measures are ordinal-based measures. Ordinal scales allow a rank order, but have unequal intervals between scores and provide nonlinear results [16]. Consequently, when using sum scores from multiple items the whole

may not equal the sum of the parts [36]. Ordinal scores are not suited for parametric statistical testing and conclusions from clinical studies based on ordinal-based outcome measures should be interpreted with great caution [37]. The Rasch-built FSHD-RODS provides an interval scale, indicating that the differences between points on the scale are measurable and equal, which enables parametric testing and comparison of changes throughout the scale [15].

Another shortcoming of most existing outcome measures is that they are not capable of measuring the entire disease spectrum and/or cannot properly differentiate between patients with different levels of functioning. It was shown that this is the case for the MFM [12–14], and similar results have been described for ordinal outcomes measures developed through the classical test theory in other neuromuscular disorders, like the Muscular Impairment Rating Scale (MIRS) staging for myotonic dystrophy type 1 [27,38].

Over the past years, drug regulatory agencies such as the FDA (United States) and the EMA (European Union) have emphasized their preference for the use of patient-relevant and preferably patient-reported outcome measures as primary outcomes for phase III clinical trials [39,40]. FSHD comprises more than just muscle weakness and outcome measures should not be focused on measures of impairment only. Multiple studies have shown that factors such as mobility, pain, physical activity level, the ability to carry out activities, and social participation are important determinants of the disease burden in FSHD and other neuromuscular diseases [41–45]. Because an improvement or decline in muscle strength has no intrinsic meaning in the absence of correlation to the disease burden or quality of life, we chose to develop a more clinically meaningful outcome measure focused on activities and participation. A next step in the clinimetric evaluation of the FSHD-RODS that is currently being undertaken is to assess its responsiveness in longitudinal studies. However, how responsiveness should be assessed should be predefined by the scientific FSHD community. Two commonly used options include relying on *p*-value-driven responsiveness or on a predefined arbitrarily chosen cut-off for the scale of interest. The first option, in particular, disregards whether the difference found

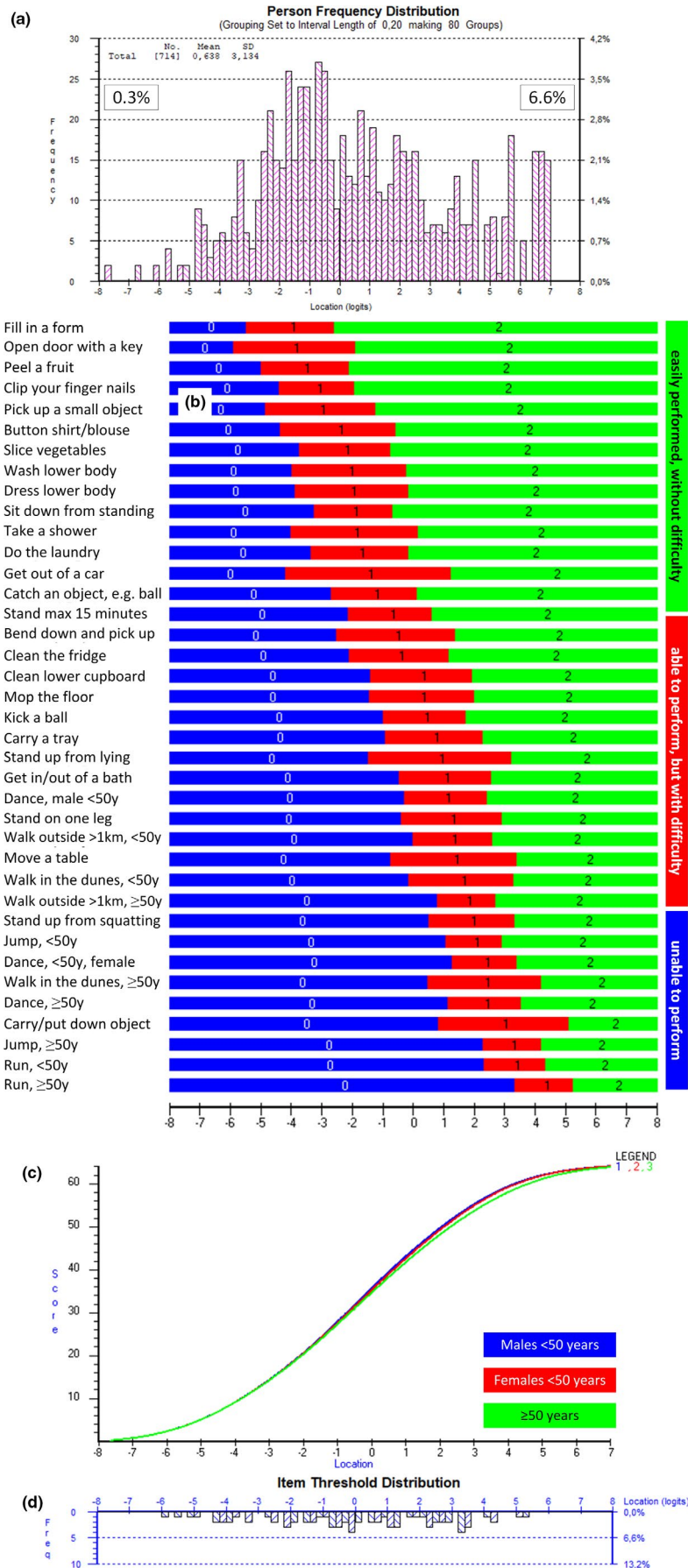


FIGURE 2 Facioscapulothoracic muscular dystrophy specific Rasch-built overall disability scale (FSHD-RODS) with 32 items. (a) Distribution of activity and participation assessment (ability location) of 714 patients with FSHD, assessed with the final FSHD-RODS. A total of 0.3% of the patients demonstrated a floor effect and 6.6% demonstrated a ceiling effect (maximum scores). (b) Threshold map of the final 32 items (38 inquiries) as part of the FSHD-RODS. The map shows the expected response for each item related to the ability of the patients using FSHD-RODS. The easiest item was "able to fill in a form", the most difficult item was "able to run" for males 50 years and older. Zero logit is set as the average of item difficulty and patient ability. This means that a patient with a mean score would be able to clean the refrigerator (this item requires -0.457 logits) easily and would have a higher probability of being able to perform the easier tasks (these having a lower logit location score); conversely, this patient will have a higher chance of experiencing extra difficulty with the more difficult tasks and will most probably fail on these. (c) Graph demonstrating the relationship between the overall raw sum scores (vertical axis: ranging from 0 to 64: 32 items, maximum score per item is $2: 32 \times 2 = 64$) with the Rasch-obtained corresponding interval scores (in logits; horizontal axis) showing the typical S-shape pattern. In essence, this graph shows the transformation of raw ordinal-based scores to interval Rasch-based values. Three S-shape figures are being presented with minor differences. (d) Graph showing the location of the 64 thresholds in the final FSHD-RODS (32 items, three response options, meaning two thresholds per item) [Colour figure can be viewed at wileyonlinelibrary.com]

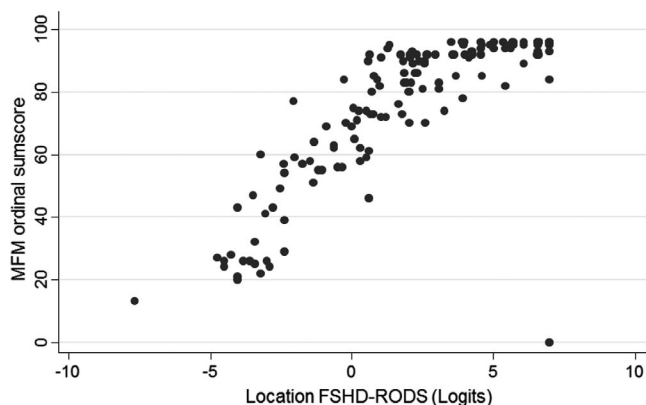


FIGURE 3 Association between the final facioscapulohumeral muscular dystrophy Rasch-built overall disability scale (FSHD-RODS) and the motor function measure (MFM) score. Significant associations were obtained between the two outcome measures (Spearman correlation coefficient 0.86)

has any clinical relevance. Therefore, a third option may be preferred that is based on the concept of minimum clinically important differences, taking into account the varying personally obtained standard errors through Rasch analyses, as has been done for other Rasch-built scales [38,46,47].

There are some limitations to this study that should be addressed. In the final Rasch-built interval scale three items had significant residual correlations. Leaving these items in resulted in a better distribution of items and corresponding thresholds. We also tested the model without these items and it still fulfilled the Rasch model expectations, but with a less optimal threshold spectrum and therefore we decided to keep these items in. Cross-cultural validation was tested for five different countries. The FSHD-RODS should

be used with caution in other countries and cultures until further cross-cultural validation has been performed [48,49].

Finally, the FSHD research community should aim to standardize the use of various outcome measures at the different levels of assessing outcome [50]. Additionally, the definition used for assessment of responsiveness could be determined. Lessons could be learned from previous efforts taken by other study groups in various neuromuscular disorders [46,51].

The FSHD-RODS is a patient-reported disease-specific interval measure suitable for detecting activity and participation restrictions in patients with FSHD. Longitudinal studies are currently ongoing to determine the functional deterioration slope in FSHD as preparation for the upcoming clinical intervention trials in FSHD.

ACKNOWLEDGMENTS

We would like to sincerely thank the UK FSHD Patient Registry for their invaluable contribution to the data collection for this study. This work was supported by the Prinses Beatrix Spierfonds.

CONFLICT OF INTEREST

The authors report no conflicts of interest relevant to the current research.

AUTHOR CONTRIBUTIONS

Karlien Mul: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Project administration (equal); Writing – original draft (equal); Writing – review and editing (equal). **Tatiana Hamadeh:** Data curation (equal); Formal analysis (equal). **Corinne Horlings:** Conceptualization (equal); Writing – review and editing (equal). **Rabi Tawil:** Data curation (equal); Project administration (equal); Writing – review and editing

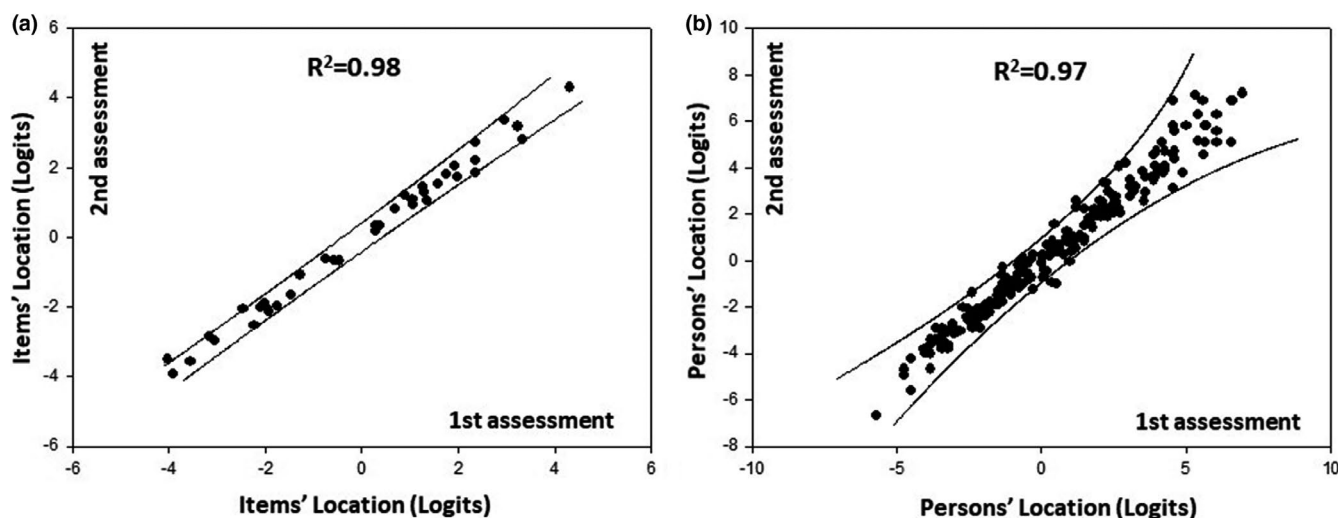


FIGURE 4 Item difficulty hierarchy of the final facioscapulohumeral muscular dystrophy Rasch-built overall disability scale (FSHD-RODS) in the first versus the second assessment. Almost all items and patients were located within the 95% confidence interval (solid lines), reflecting ideal reliability

		Unable to perform	Able to perform, but with difficulty	Easily performed, without difficulty
Are you able to:		0	1	2
1	Fill in a form			
2	Open a door with a key			
3	Peel an apple/orange			
4	Clip your finger nails			
5	Pick up a small object			
6	Button a shirt/blouse			
7	Slice vegetables			
8	Wash your lower body			
9	Dress your lower body			
10	Sit down from a standing position			
11	Take a shower			
12	Put laundry in the washing machine			
13	Get out of a car			
14	Catch an object, e.g. a ball			
15	Remain standing for a short period of time, e.g. max 15 minutes			
16	Bend forward and pick something up			
17	Clean the fridge			
18	Clean the bottom of a kitchen cupboard			
19	Mop a floor			
20	Kick a ball			
21	Carry a tray			
22	Stand up from lying down			
23	Get in or out of a bath			
24	Stand on one leg			
25	Move a table			
26	Stand up from a squatting position			
27	Carry and put down a heavy object			
28	Dance			
29	Walk through the dunes			
30	Walk outdoors, for more than 0.6 mile (1 km)			
31	Jump			
32	Run			

TABLE 2 Final 32-item facioscapulohumeral muscular dystrophy specific Rasch-built overall disability scale

Note: The final 32-item facioscapulohumeral muscular dystrophy specific Rasch-built overall disability scale (FSHD-RODS) is presented that fulfilled all Rasch model expectations. The easiest item was found to be ability to fill in a form, and the most difficult was ability to run (for 50 years and older). Note: the presented raw scores in the Table should be transformed to interval scores (available on request) for proper application, and permission to use this scale should be obtained through the authors. An instruction manual is also available to standardize and promote proper interpretation of the items (also available on request).

(equal). **Jeffrey Statland**: Data curation (equal); Project administration (equal); Writing – review and editing (equal). **Sabrina Sacconi**: Data curation (equal); Project administration (equal); Writing – review and editing (equal). **Alastair Corbett**: Data curation (equal); Project administration (equal); Writing – review and editing (equal). **Nicol Voermans**: Writing – review and editing (equal). **Catharina G. Faber**: Conceptualization (equal); Methodology (equal); Writing – review and editing (equal). **B G van Engelen**: Conceptualization (equal); Funding acquisition (equal); Supervision (equal); Writing – review and editing (equal). **Ingemar S.J. Merkies**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Project administration (equal); Supervision (equal); Validation (equal); Writing – original draft (equal); Writing – review and editing (equal).

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Karliën Mul  <https://orcid.org/0000-0001-8487-9478>

Tatiana Hamadeh  <https://orcid.org/0000-0003-2413-8218>

REFERENCES

- Mul K, Lassche S, Voermans NC, et al. What's in a name? The clinical features of facioscapulohumeral muscular dystrophy. *Pract Neurol*. 2016;16(3):201-207.
- Bouwman LF, van der Maarel SM, de Greef JC. The prospects of targeting DUX4 in facioscapulohumeral muscular dystrophy. *Curr Opin Neurol*. 2020;33(5):635-640.
- LoRusso S, Johnson NE, McDermott MP, et al. Clinical trial readiness to solve barriers to drug development in FSHD (ReSolve): protocol of a large, international, multi-center prospective study. *BMC Neurol*. 2019;19(1):224.
- Le Gall L, Sidlauskaitė E, Mariot V, Dumonceaux J. Therapeutic strategies targeting DUX4 in FSHD. *J Clin Med*. 2020;9(9):2886.
- Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. *Clin Pharmacol Ther*. 2001;69(3):89-95.
- Tawil R, Shaw DW, van der Maarel SM, Tapscott SJ. Clinical trial preparedness in facioscapulohumeral dystrophy: outcome measures and patient access: 8–9 April 2013, Leiden, The Netherlands. *Neuromuscul Disord*. 2014;24(1):79-85.
- U.S. Department of Health and Human Services FaDA. Guidance for Industry. Patient-reported outcome measures: Use in Medical Product Development to Support Labeling Claims; 2009.
- EMA. *Appendix 2 to the guideline on the evaluation of anticancer medicinal products in man. The use of patient-reported outcome (PRO) measures in oncology studies*, Amsterdam: EMA; 2016.
- Mul K, Kinoshita J, Dawkins H, et al. 225th ENMC international workshop: A global FSHD registry framework, 18–20 November 2016, Heemskerk, The Netherlands. *Neuromuscul Disord*. 2017;27(8):782-790.
- Tawil R, Padberg GW, Shaw DW, et al. Clinical trial preparedness in facioscapulohumeral muscular dystrophy: clinical, tissue, and imaging outcome measures 29–30 May 2015, Rochester, New York. *Neuromuscul Disord*. 2016;26(2):181-186.
- Eichinger K, Heatwole C, Iyadurai S, et al. Facioscapulohumeral muscular dystrophy functional composite outcome measure. *Muscle Nerve*. 2018;58:72-78.
- de Lattre C, Rippert P, Hmaroun D, Sacconi S, Poirot I, Vuillerot C. The motor function measure (MFM) in the facio scapulo humeral dystrophy (FSHD) population: description and responsiveness. *Ann Phys Rehabil Med*. 2016;59S:e84-e85.
- Guillot T, Roche S, Rippert P, et al. Is going beyond Rasch analysis necessary to assess the construct validity of a motor function scale? *Arch Phys Med Rehabil*. 2018;99(9):1776-1782.e1779.
- Mul K, Horlings CGC, Faber CG, van Engelen BGM, Merkies ISJ. Rasch analysis to evaluate the Motor Function Measure for patients with facioscapulohumeral muscular dystrophy. *Int J Rehabil Res*. 2020;44:38-44.
- Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. *Arch Phys Med Rehabil*. 1989;70(12):857-860.
- Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. *Arch Phys Med Rehabil*. 1989;70(4):308-312.
- Tennant A, Conaghan PG. The Rasch measurement model in rheumatology: what is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis Rheum*. 2007;57(8):1358-1362.
- Streiner DL, Norman GR. *Health measurement scales. A practical guide to their development and use*, 2nd edn. New York, NY: Oxford University Press; 1998.
- Merkies IS, Lauria G, Faber CG. Outcome measures in peripheral neuropathies: requirements through statements. *Curr Opin Neurol*. 2012;25(5):556-563.
- Mul K, Vincenten SCC, Voermans NC, et al. Adding quantitative muscle MRI to the FSHD clinical trial toolbox. *Neurology*. 2017;89(20):2057-2065.
- Mul K, Voermans NC, Lemmers R, et al. Phenotype-genotype relations in facioscapulohumeral muscular dystrophy type 1. *Clin Genet*. 2018;94(6):521-527.
- Evangelista T, Wood L, Fernandez-Torron R, et al. Design, set-up and utility of the UK facioscapulohumeral muscular dystrophy patient registry. *J Neurol*. 2016;263(7):1401-1408.
- Guien C, Blandin G, Lahaut P, et al. The French National Registry of patients with Facioscapulohumeral muscular dystrophy. *Orphanet J Rare Dis*. 2018;13(1):218.
- Hilbert JE, Kissel JT, Luebke EA, et al. If you build a rare disease registry, will they enroll and will they use it? Methods and data from the National Registry of Myotonic Dystrophy (DM) and Facioscapulohumeral Muscular Dystrophy (FSHD). *Contemp Clin Trials*. 2012;33(2):302-311.
- van Nes SI, Vanhoutte EK, van Doorn PA, et al. Rasch-built Overall Disability Scale (R-ODS) for immune-mediated peripheral neuropathies. *Neurology*. 2011;76(4):337-345.
- WHO. *The International Classification of Functioning, Disability and Health*. Geneva: World Health Organization; 2001.
- Hermans MC, Faber CG, De Baets MH, et al. Rasch-built myotonic dystrophy type 1 activity and participation scale (DM1-Activ). *Neuromuscul Disord*. 2010;20(5):310-318.
- Berard C, Payan C, Hodgkinson I, et al. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord*. 2005;15(7):463-470.
- Andrich D, Sheridan B, Luo G. Rasch models for measurement: RUMM2030. Perth; 2010.
- Rasch G. *Probabilistic models for some intelligence and attainment tests*. Chicago: University of Chicago Press; 1960.
- Pallant JF, Tennant A. An introduction to the Rasch measurement model: an example using the Hospital Anxiety and Depression Scale (HADS). *Br J Clin Psychol*. 2007;46(Pt 1):1-18.
- Fisher WP. Reliability statistics. *Rasch Meas Transact*. 1992;6:238.
- Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statistics and principal component analysis of residuals. *J Appl Meas*. 2002;3(2):205-231.

34. Wright BD, Stone MH. *Best test design*. Chicago: Media Press; 1979.
35. Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ*. 1995;310(6973):170.
36. Stucki G, Daltroy L, Katz JN, et al. Interpretation of change scores in ordinal clinical scales and health status measures: the whole may not equal the sum of the parts. *J Clin Epidemiol*. 1996;49(7):711-717.
37. Grimby G, Tennant A, Tesio L. The use of raw scores from ordinal scales: time to end malpractice? *J Rehabil Med*. 2012;44(2):97-98.
38. Draak TH, Vanhoutte EK, van Nes SI, et al. Changing outcome in inflammatory neuropathies: Rasch-comparative responsiveness. *Neurology*. 2014;83(23):2124-2132.
39. Mendell JR, Csimma C, McDonald CM, et al. Challenges in drug development for muscle disease: a stakeholders' meeting. *Muscle Nerve*. 2007;35(1):8-16.
40. Tawil R, van der Maarel S, Padberg GW, van Engelen BG. 171st ENMC international workshop: Standards of care and management of facioscapulohumeral muscular dystrophy. *Neuromuscul Disord*. 2010;20(7):471-475.
41. Kalkman JS, Schillings ML, Zwarts MJ, et al. The development of a model of fatigue in neuromuscular disorders: a longitudinal study. *J Psychosom Res*. 2007;62(5):571-579.
42. Johnson NE, Quinn C, Eastwood E, et al. Patient-identified disease burden in facioscapulohumeral muscular dystrophy. *Muscle Nerve*. 2012;46(6):951-953.
43. Graham CD, Weinman J, Sadjadi R, et al. A multicentre postal survey investigating the contribution of illness perceptions, coping and optimism to quality of life and mood in adults with muscle disease. *Clin Rehabil*. 2014;28(5):508-519.
44. Molenaar DS, Vermeulen M, de Visser M, de Haan R. Impact of neurologic signs and symptoms on functional status in peripheral neuropathies. *Neurology*. 1999;52(1):151-156.
45. Merkies IS, Hughes RA, Donofrio P, et al. Understanding the consequences of chronic inflammatory demyelinating polyradiculoneuropathy from impairments to activity and participation restrictions and reduced quality of life: the ICE study. *J Peripher Nerv Syst*. 2010;15(3):208-215.
46. Vanhoutte EK, Faber CG, Merkies IS. 196th ENMC international workshop: Outcome measures in inflammatory peripheral neuropathies 8-10 February 2013, Naarden, The Netherlands. *Neuromuscul Disord*. 2013;23(11):924-933.
47. Hobart J, Cano S. Improving the evaluation of therapeutic interventions in multiple sclerosis: the role of new psychometric methods. *Health Technol Assess*. 2009;13(12):1-177.
48. Kucukdeveci AA, Sahin H, Ataman S, Griffiths B, Tennant A. Issues in cross-cultural validity: example from the adaptation, reliability, and validity testing of a Turkish version of the Stanford Health Assessment Questionnaire. *Arthritis Rheum*. 2004;51(1):14-19.
49. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417-1432.
50. Mul K, van den Boogaard ML, van der Maarel SM, van Engelen BG. Integrating clinical and genetic observations in facioscapulohumeral muscular dystrophy. *Curr Opin Neurol*. 2016;29(5):606-613.
51. CI-PERINOMS Study Group. CI-PERINOMS: chemotherapy-induced peripheral neuropathy outcome measures study. *J Peripher Nerv Syst*. 2009;14(2):69-71.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Mul K, Hamadeh T, Horlings CG, et al. The facioscapulohumeral muscular dystrophy Rasch-built overall disability scale (FSHD-RODS). *Eur J Neurol*. 2021;28:2339-2348. <https://doi.org/10.1111/ene.14863>