Development of a patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy: DMD Upper Limb PROM

Peer-reviewed author version

KLINGELS, Katrijn; Mayhew, Anna G; Mazzone, Elena Stacy; Duong, Tina; Decostre, Valérie; Werlauff, Ulla; Vroom, Elizabeth; Mercuri, Eugenio & Goemans, Nathalie M (2017) Development of a patient-reported outcome measure for upper limb function in Duchenne muscular dystrophy: DMD Upper Limb PROM. In: DEVELOPMENTAL MEDICINE AND CHILD NEUROLOGY, 59(2), p. 224-231.

DOI: 10.1111/dmcn.13277
Handle: http://hdl.handle.net/1942/23559

Development of a patient-reported outcome measure for upper limb function in Duchenne Muscular Dystrophy (DMD-Upper limb PROM)

Klingels K*[1,2], Mayhew AG*[3], Duong T[4], Decostre V[5], Werlauff U[6], De Groot I[7], Vroom E[8], Mercuri E[9], Goemans NM[10]
on behalf of the Upper Limb Clinical Outcome group

[1]KU Leuven, Department of Rehabilitation Sciences, Leuven, Belgium;
[2]Hasselt University, Rehabilitation Research Center (REVAL), Biomed, Diepenbeek, Belgium;
[3]John Walton Muscular Dystrophy Research Centre, Institute of Genetic Medicine, Newcastle Upon Tyne, UK;
[4]Stanford University School of Medicine, Department of Neurology, US;
[5]Institut de Myologie, GH Pitié Salpêtrière, Paris, France;
[6]Danish National Rehabilitation Center for Neuromuscular Diseases , Aarhus , Denmark; [10]Radboud
[7]University Medical Centre, Nijmegen, The Netherlands;

[8]Duchenne Parent Project, Veenendaal, The Netherlands.
[9]Department of Paediatric Neurology, Catholic University, Rome, Italy,
[10]University Hospitals Leuven, Department of Child Neurology, Leuven, Belgium;

** both first authors*

Word count 2999 – revised version 3111

Corresponding author

Dr. Katrijn Klingels

Tervuursevest 101

3000 Leuven

Belgium

katrijn.klingels@kuleuven.be

Abstract

Aim

To develop a patient-reported outcome measure (PROM) assessing upper limb function related to activities of daily living (ADL) that cannot be observed in a clinical setting, specifically for patients with Duchenne Muscular Dystrophy (DMD) across a wide age range and applicable in the different disease stages.

Methods

The developmental process was based on US Food and Drug Administration guidelines. This included item generation from a systematic review of existing tools and expert opinion on task difficulty and relevance, involving individuals with DMD. Cultural aspects affecting ADL were taken into consideration to make this tool applicable to the broad DMD community. Items were selected in relation to a conceptual framework reflecting disease progression covering the full range of upper limb function across different ADL domains.

Results

After pilot testing and iterative Rasch analyses, redundant or clinically irrelevant items were removed. The final questionnaire consists of 32 items covering four domains of ADL (Food, Self-care, Household and environment, Leisure and communication). Test-retest reliability was excellent.

Interpretation

A DMD specific Upper Limb PROM was developed based on clinical relevance and psychometric robustness. The main purpose is to document the patient self-reported natural history of DMD and assess the efficacy of interventions.

Short title

DMD-Upper Limb PROM

What this paper adds

- A new patient-reported outcome measure for DMD was developed, targeting upper limb function in daily life.
- Modern psychometric techniques confirmed its uni-dimensionality, internal consistency and test-retest reliability.
- Involvement of different stakeholders, including young men with DMD in the developmental process guaranteed the clinical relevance of the tool.

The emergence of new therapeutic strategies for Duchenne Muscular Dystrophy (DMD) has exposed the need for suitable outcome measures to assess their efficacy within the framework of clinical trials. So far, the focus of trials has been put on the ambulant stage of the disease. Extensive research has been done on the clinical feasibility and psychometric properties of the six-minute walk test (6MWT), North Star Ambulatory Assessment (NSAA) and the timed function tests.[1-3] However, with longer trials and post marketing requirements there is a need for outcome measures encompassing different stages of the disease.

The next step therefore is to invest in developing outcome measures describing disease progression in the upper limbs from early ambulant stages over transition stages to non-ambulant stages. This would allow a better understanding of disease evolution and efficacy of interventions throughout the life span.

Upper limb weakness manifests when the boys are still ambulant and is seen first in the proximal muscles.[4-5] Further progression of weakness is evident in a proximal to distal gradient, and upper limb function becomes more and more influenced by weakness in conjunction with contractures and/or growth resulting in compensatory strategies and ultimately loss of function. In the end stage, movements are limited to the fingers. Muscle weakness impacts activities of upper limb function and performance of activities in daily life (ADL). The relationship between muscle strength and function is not linear but influenced by compensatory behavior, personal (such as age, lifestyle, motivation) and environmental factors. Therefore, the assessment of limitations in upper limb function in daily life warrants further investigation.

A multidisciplinary international Clinical Outcomes Group highlighted the lack of suitable outcome measures within this context,[6] and as a first step they developed the Performance of the Upper Limb module (PUL), a Performance Outcome measure (PerfO) specifically designed for upper limb function in patients with DMD.[7] Reliability and responsiveness to change for this tool have been published.[8] In addition to the PUL, this group also identified the need to develop in parallel a patient reported outcome measure (PROM) to complement information on ADL which cannot otherwise be observed in a clinical or research setting, for example feeding, washing, and leisure activities and which focuses on outcomes that are clinically meaningful in the patients daily life and may impact on their quality of life.

Therefore, the aim was to develop a PROM on upper limb function in ADL specifically for patients with DMD across a wide age range and applicable in the different stages of the disease. Such an instrument provides a measure of the patients' level of independence and ability to interact with the environment. The main purpose of the questionnaire is to be used for description of natural history in DMD and to assess the efficacy of interventions in daily life.

Methods

QUESTIONNAIRE DEVELOPMENT

The development of the questionnaire was based on PROM guidelines published by the US Food and Drug Administration,[9] and followed a number of steps (Figure 1). Five face-to-face meetings were held in Rome (Italy), Leuven (Belgium), Amsterdam (the Netherlands) between 2012 and 2015, organized by Duchenne Parent Project. Focus groups involved medical doctors, researchers, physiotherapists and clinicians working with DMD patients as well as representatives from patients, advocacy groups and industries.

Development of a conceptual framework

The concept to be measured was defined as upper limb function within the activity domain of the International Classification of Functioning, Disability and Health (ICF). In relation with the PUL,[7] the questionnaire should reflect the progression of weakness and the natural history of functional decline in daily life of DMD. Items should be selected to cover the widest range of upper limb activities in daily life from unimanual to bimanual activities and dexterity.

Systematic and critical review of existing questionnaires

A systematic review was performed to identify existing questionnaires used in DMD assessing aspects related to upper limb function in daily life. Five questionnaires were identified, the Egen Klassifikation,[10] the ACTIVLIM,[11] the ABILHAND,[12] the Muscular Dystrophy Functional Rating Scale (MDFRS)[13], and the Duchenne muscular dystrophy Functional Ability Self-Assessment Tool (DMDSAT).[14]. All these questionnaires assessed a wide range of abilities whereas the ABILHAND has specifically been developed for assessing manual abilities. Details of the results of this review have been reported.[15] However, none of these questionnaires covers the whole spectrum of abilities found in ambulant and non-ambulant DMD boys. Therefore, all items were submitted to the multidisciplinary international Clinical Outcomes group with the aim to determine their clinical meaningfulness for the target group as well as to propose other relevant items not included in the original item set.

Selection and adaptation of items

Based on clinical relevance and potential to measure upper limb function in daily life, items were selected by the working group. Individuals with DMD were involved throughout the iterative process by offering their opinion on task difficulty and relevance. Also, cultural aspects were taken into consideration to make this tool applicable to a broad international DMD community. Items were selected in relation to the different domains of the PUL from antigravity shoulder movements to limited finger movements in order to cover the full range of upper limb function. Items were further refined, added, or eliminated based on feedback from multidisciplinary group and individuals with DMD. The scoring was also critically reviewed and adapted to reflect disease progression and reduce possible floor and ceiling effects. A pilot pro forma was developed and tested in an international multi-centric setting. The pro forma included 39 items categorized in four domains of daily life: food, household, self-care and leisure/communication.

The PROM was officially translated in French, Dutch, German, Italian, Portuguese, Spanish, Turkish, Danish and Swedish. This process involved translation by official translators, which was corrected afterwards by a native speaking expert working in the field of neuromuscular disorders, followed by backward translation to guarantee the accuracy of the translation. The questionnaires were self-completed by the boys, when possible or by their parents or caregivers.

ANALYSIS REVIEW

Rasch model

The Rasch model estimates the item difficulty and a participant's ability on a common linear scale based on the responses provided for each item. The model assumes that participants with a higher ability level have a higher probability to positively respond to an item compared to participants with a lower ability.Every participant's ability and each item difficulty are represented by a score, expressed in logits (log odds units), a linear unit defined as the natural logarithm of the odds of positive achievement of any item by a participant. The Rasch analysis was computed with RUMM

2030 version 5.1 for Windows. This program provides fit statistics, which expresses the extent to which each item contributes to a single dimension (upper limb function in daily life). The statistics include overall fit statistics (item trait interaction and person separation index), and fit statistics for individual items and persons. The item–trait interaction (a chi-square value) reflects the degree of invariance across the aimed dimension. The probability exceeding 0.05 indicates that the data fit the model. The person separation index (PSI) is an estimate of the internal consistency reliability and is similar to chronbach's alpha. The individual item-fit and person-fit statistics are represented by their item locations (in logits), standard errors, residuals and fit to the model. The residual values express a source of misfit to the model, and are defined as being greater than +-2.5.[16-18]

For test-retest reliability, 15 boys with DMD of one centre (Leuven) or their caregivers completed the questionnaire twice with a time interval between 7 and 14 days. This interval was chosen as it is long enough not to remember the responses, but short enough to avoid actual changes in the ability level due to the progression of the disease. Intra-class correlation coefficient (2.1) was calculated.

Discriminative ability of the DMD-Upper limb PROM, which would confirm construct validity, was examined by testing the differences in total scores between the different Brooke levels[19] with one-way Analysis of Variance (ANOVA) and post hoc Bonferroni t-tests, using SAS Enterprise guide.

PARTICIPANTS

Participants were recruited between 2013 and 2015 through eight international specialized centers (Rome, Italy; London, England; Leuven, Belgium; Messina, Italy; Newcastle, England; Paris, France; Nijmegen, The Netherlands; Aarhus, Denmark). Subjects were eligible if they were male, had a diagnosis of DMD and were aged 7 years or above. All patients or their parents gave informed consent to participation.

RESULTS

The first experimental version of the PROM included 39 items. For each question, the boys and their parents were asked to fill in their perceived difficulty to perform the activity on a four-level ordinal scale (cannot do – very hard – a little hard – easy). A Rasch analysis was performed on the results of 101 assessments (mean age 12 years, SD 4 years; ambulant N=52, non-ambulant N=49). In response to this primary analysis, six items were deleted because of issues with item fit, disordered thresholds, issues of dependency and targeting (ceiling effect).

The second experimental version included 33 items and four response categories (cannot do – very hard – a little hard – easy). A second Rasch analysis was performed in 357 assessments (mean age 12 years, SD 4 years, ambulant N=198, non-ambulant N=159). In response to this second analysis, one item was deleted (*use touch pad*) due to redundancy and the number of response categories was reduced from four to three because of disordered thresholds. Feedback from patient groups and clinicians suggested that a lack of distinction between the categories of "very hard" and "a little hard" could be a reason for lack of ordered thresholds for many of the items.

The final questionnaire consisted of 32 items covering four domains of ADL (1) Food: 7 items, 2) Self-care: 8 items, 3) Household and environment: 6 items, 4) Leisure and communication: 11 items). For each question, the boys and their parents were asked to fill in their perceived difficulty to perform the activity on a three-level scale (cannot do - can do with difficulty - can do easily). The time to administer the PROM never exceeded 10 minutes. . We recommend using the questionnaire in boys from 7 years of age and having it completed by parents or caregivers up to the age of 16 years.

A total of 194 subjects with DMD recruited across eight international specialized centers for DMD completed the third and final version of the PROM questionnaire. Subjects had a mean age of 15 years (SD 7 years, range 7-43 years). Eighty-four subjects (43%) were ambulant and 128 subjects (66%) were on steroids. The questionnaire was completed by caregivers in 130 subjects (67%) and 64 boys (23%) independently completed the questionnaire.

Table I presents the item fit ordered by difficulty level in terms of functional ability. No items displayed mode misfit in terms of estimated residuals but *dial on a cell phone* and *write several lines* had a significant chi-square probability, suggesting that these items do not discriminate well. The item locations spread out well (-4.49 to 3.79), indicating that the PROM defines a good continuum with little overlap (i.e. very few items measure the same level of ability).

Fit of individual responses was adequate (persons mean fit residual: -0.35, SD: 0.68). Subjects were evenly distributed across ability levels (Figure 2). The range of the upper histogram horizontally that extends outside the range of the lower histogram suggests acceptable floor and ceiling levels (in 11 and 5 children, respectively).

The item response option thresholds were ordered for 30 out of 32 items, indicating that the proposed scoring system worked as intended for most items (Figure 3). Disordered thresholds were found in two items (*pick up pen* and *pull up trousers*) indicating that subjects had difficulty discriminating between response categories given their functional ability. Category probability curves showed that in these two items the category 'with difficulty' did not emerge as the most likely to be selected at any point upon the underlying scale, meaning that the perception of these items is more binary in nature, and these activities are perceived as either impossible, or easy, with very rare intermediate responses.

The overall item-trait interaction chi-square value was 174.33 (64 degrees of freedom ($p<0.0001$). Uni-dimensionality was acceptable (t-test 9.2%, binomial test lower 95% confidence interval proportion, 0.05). The person separation index (PSI) was estimated at 0.97, indicating very high reliability.

Only two items (*Press buttons on elevator* and *Turn a light switch on*) displayed a positive correlation above 0.40 ($r=0.41$). These items would be expected to show that even though the tasks are different they require a similar level of ability to perform. This dependency could artificially inflate reliability as measured by the PSI. However, removing the items from analysis did not significantly alter this index (PSI=0.97). As they are both important from a clinical point of view, it was decided to keep both items in the questionnaire.

For discriminant ability, significant differences were found between the Brooke levels ($F=248.15$; $p<0.0001$). Figure 4 shows decreasing total PROM scores (expressed in percentage) with increasing Brooke levels. Post hoc comparisons yielded significant differences between the levels, except for comparisons between adjacent levels 1-2, 3-4 and 5-6.

Test-retest reliability

Intra-class correlation coefficient of 0.99 showed excellent agreement between the first and second assessment.

Discussion

The aim of this study was to develop a DMD specific patient reported outcome measure (PROM) and evaluate its ability to measure the underlying construct of upper limb function in daily life in boys and young men with DMD. The clinical relevance of the questionnaire was guaranteed by involving all stakeholders including the patient community. Its internal reliability and validity has been demonstrated based on modern psychometric methods in a large set of patients from eight different centres.

New therapy developments for DMD have highlighted the need for disease specific instruments *assessing* activity limitations in the context of clinical trials. An extensive literature search identified a gap in suitable outcome measures for the assessment of upper limb function in DMD across all stages of the disease.[15] A multidisciplinary international Clinical Outcomes group consisting of clinicians, scientists, industries, young men with DMD, and patient advocacy groups developed an observer rated tool, specifically designed for assessing upper limb function in ambulant and non-ambulant DMD (PUL).[7] However, there was still a need to better understand patient and caregiver perception of abilities related to upper limb function in daily life, which cannot otherwise be observed in a clinical or research setting, for example feeding, washing, and leisure activities.

The final 32 items in this questionnaire cover four domains of ADL (Food, Self-care, Household and environment, Leisure and communication). They were selected based on clinical relevance and iterative Rasch analyses and showed a good fit with nearly all the response categories working appropriately. The simplified scoring of four to three categories contributed to this improvement. These findings suggest that this scale is suitable for assessing perception of upper limb function in daily life in this population. The first and second experimental version testing included only few adult patients and to investigate the possible floor effect the age range was extended in the final testing. The targeting of the scale showed minimal issues with floor and ceiling effect, however these issues may well be part of the population's natural course as there is a point when upper limb function is fully preserved in either weaker ambulant boys or in stronger non-ambulant boys. Also men in the later stages of the disease may reach a point where they have no useful upper limb function due to progressive contractures and muscle weakness. Eleven patients aged between 23 and 34 years old with Brooke levels 5 or 6 had the minimum score on the DMD-Upper limb PROM. These adults were in such an advanced disease stage that they either had no upper limb function left (Brooke level 6) or only had the ability to drive their electric wheelchair with joystick control that requires minimal hand movement (Brooke level 5).

The strength of this questionnaire compared to other existing scales assessing activities in daily life is that 1) it focuses on upper limb function in daily life (uni-dimensional) and 2) it is DMD specific, covering the whole spectrum of disease evolution.

This questionnaire has been developed for young men/boys with DMD from the age of 7 years, an age at which a maximum score on the PROM may be obtained in typically developing children. We recommend having it completed by parents or caregivers up to the age of 16 years as previous publications have indicated that parents more accurately discriminate function in a likert-type rating scale compared to children who tend to perceive activities in a dichotomous way ("impossible" or "easy").[20] However, inherent in questionnaires completed by parents or caregivers is the "proxy effect", which means inaccurate reporting about characteristics of others, which poses a threat to the validity of information.[21] Therefore, it was recommended to ask the individual to self report the questionnaire from the age of 17 years.

We acknowledge several limitations in our study. Firstly, our data set was formed from cross-sectional data. Despite the fact that our data provided a sufficient sample, longitudinal data analysis is needed to assess stability of the scale over time and its sensitivity to capture disease progression in

a population where change appears to occur over a long period of time. Secondly, this study was unable to assess construct validity given that no gold standard in functional assessment for this population exists. Nevertheless, the development of a conceptual framework reflecting expert opinion on disease progression in upper limb function in DMD, supports the construct to be measured by this questionnaire. Finally, an inherent challenge in the development of an internationally sensitive PROM assessing function is the inability to control for external impact factors such as socio-economic and cultural aspects, as well as environmental adaptations that would intervene with the answers of the respondents. Therefore, experts from different centers in Europe and in US were involved to ensure large applicability of this scale independent of cultural and environmental factors. Further consideration for Africa and Asia should be taken in the coming steps.

In conclusion, this collaborative effort resulted in the creation of the DMD-Upper Limb PROM.. Applied modern psychometric methods in conjunction with specialist clinical knowledge, patient and family feedback has ensured that the DMD-Upper Limb PROM is not only disease and population specific but also assesses clinically meaningful upper limb function in this population. The main purpose of this questionnaire is to describe patient self-reported natural history in DMD and to assess the efficacy of interventions in daily life. Although the metric properties allow capturing changes in upper limb function in daily life over time and following intervention, further study is warranted on the responsiveness and predictiveness of the DMD-Upper limb PROM.

Acknowledgement

Other participants in the Upper Limb Clinical Outcome group were: Mazzone ES, Italy; Eagle M, England; Main M, England; Messina S, Italy; Campion G, The Netherlands; Servais L, Paris; Van den Hauwe M,Es E, Belgium; Pane P, Rome; Buccella F, DPP Italy; Kuijer J, The Netherlands, Ceradini F, DPP Italy; Henrikson E, US; McDonald C, US; Florence J, US.

The study is supported by Duchenne Parent Project, the Netherlands. The authors thank the patients for their contribution and Duchenne Parent Project, Italy for organizing and coordinating the workshops.

References

1. Mazzone E, Martinelli D, Berardinelli A, et al. North Star Ambulatory Assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscul Disord* 2010; **20**: 712-6.
2. Mayhew A, Cano S, Scott E, Eagle M, Bushby K, Muntoni F; North Star Clinical Network for Paediatric Neuromuscular Disease. Moving towards meaningful measurement: Rasch analysis of the North Star Ambulatory Assessment in Duchenne muscular dystrophy. *Dev Med Child Neurol* 2011; **53**: 535-42.
3. McDonald CM, Henricson EK, Han JJ, et al. The 6-minute walk test in Duchenne/Becker muscular dystrophy: longitudinal observations. *Muscle Nerve* 2010; **42**: 966-74.
4. Seferian AM, Moraux A, Annoussamy M, et al. Upper limb strength and function changes during a one-year follow-up in non-ambulant patients with Duchenne Muscular Dystrophy: an observational multicenter trial. *PLoS One* 2015; **10**: e0113999.
5. Janssen MM, Bergsma A, Geurts AC, de Groot IJ. Patterns of decline in upper limb function of boys and men with DMD: an international survey. J Neurol 2014; 261:1 269-88.
6. Mercuri E, McDonald C, Mayhew A, et al. International workshop on assessment of upper limb function in Duchenne Muscular Dystrophy: Rome, 15-16 February 2012. *Neuromuscul Disord* 2012; **22**: 1025-8.
7. Mayhew A, Mazzone ES, Eagle M, et al. Performance of the Upper Limb Working Group. Development of the Performance of the Upper Limb module for Duchenne muscular dystrophy. *Dev Med Child Neurol* 2013; **55**: 1038-45.
8. Pane M, Mazzone ES, Fanelli L, et al. Reliability of the Performance of Upper Limb assessment in Duchenne muscular dystrophy. *Neuromuscul Disord* 2014; **24**: 201-6.
9. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf
10. Steffensen B, Hyde S, Lyager S, Mattsson E. Validity of the EK scale: a functional assessment of non-ambulatory individuals with Duchenne muscular dystrophy or spinal muscular atrophy. *Physiother Res Int* 2001; **63**: 119-34.
11. Vandervelde L, Van den Bergh PY, Goemans N, Thonnard JL. ACTIVLIM: a Rasch-built measure of activity limitations in children and adults with neuromuscular disorders. *Neuromuscul Disord* 2007; **17**: 459–69.
12. Vandervelde L, Van den Bergh PY, Penta M, Thonnard JL. Validation of the ABILHAND questionnaire to measure manual ability in children and adults with neuromuscular disorders. *J Neurol Neurosurg Psychiatry* 2010; **81**: 506–12.
13. Lue YJ, Su CY, Yang RC, et al. Development and validation of a muscular dystrophy-specific functional rating scale. *Clin Rehabil* 2006; **20**: 804–17.
14. Landfeldt E, Mayhew A, Eagle M, et al. Development and psychometric analysis of the Duchenne muscular dystrophy Functional Ability Self-Assessment Tool (DMDSAT). *Neuromuscul Disord* 2015; **25**: 937-44.
15. Mazzone ES, Vasco G, Palermo C, et al. A critical review of functional assessment tools for upper limbs in Duchenne muscular dystrophy. *Dev Med Child Neurol* 2012; **54**: 879-85.
16. Andrich D. Rasch Models for Measurement. Beverly Hills, CA: Sage Publications; 1988.

17. Tennant A, McKenna SP, Hagell P. Application of Rasch analysis in the development and application of quality of life instruments. *Value Health* 2004; **7**:S22-6.
18. Van de Winkel A, Feys H, Lincoln N, De Weerdt W. Assessment of arm function in stroke patients: Rivermed Motor Assessment arm section revised with Rasch analysis. *Clinical Rehabilitation* 2004; 21: 471-9.
19. Brooke MH, Griggs RC, Mendell JR, et al. Clinical trial in Duchenne dystrophy. I. The design of the protocol. *Muscle Nerve* 1981; 4: 186–97.
20. Chambers CT, Johnston C. Developmental differences in children's use of rating scales. *J Pediatr Psychol* 2002; **27**: 27–36.
21. Tennant A, Badley EM, Sullivan M. Investigating the proxy effect and the saliency principle in household based postal questionnaires. *J Epidemiol Community Health* 1991; **45**:312-6.

Figure 1: Flowchart of the development process of the Upper Limb patient reported outcome measure for DMD boys (DMD-Upper Limb PROM)

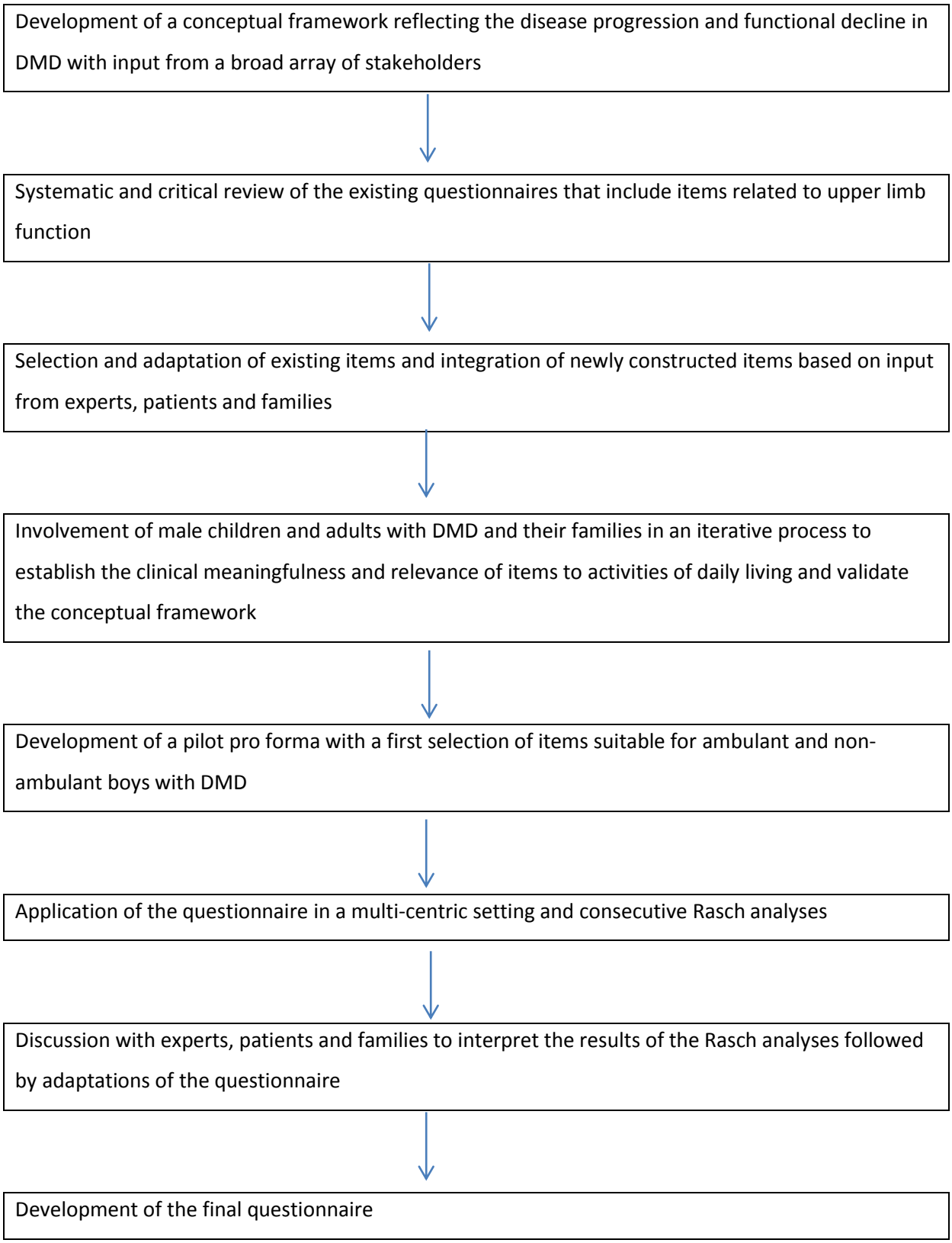


Figure 2: Person-item location distribution. Targeting of the patient sample (top) to the items (bottom)

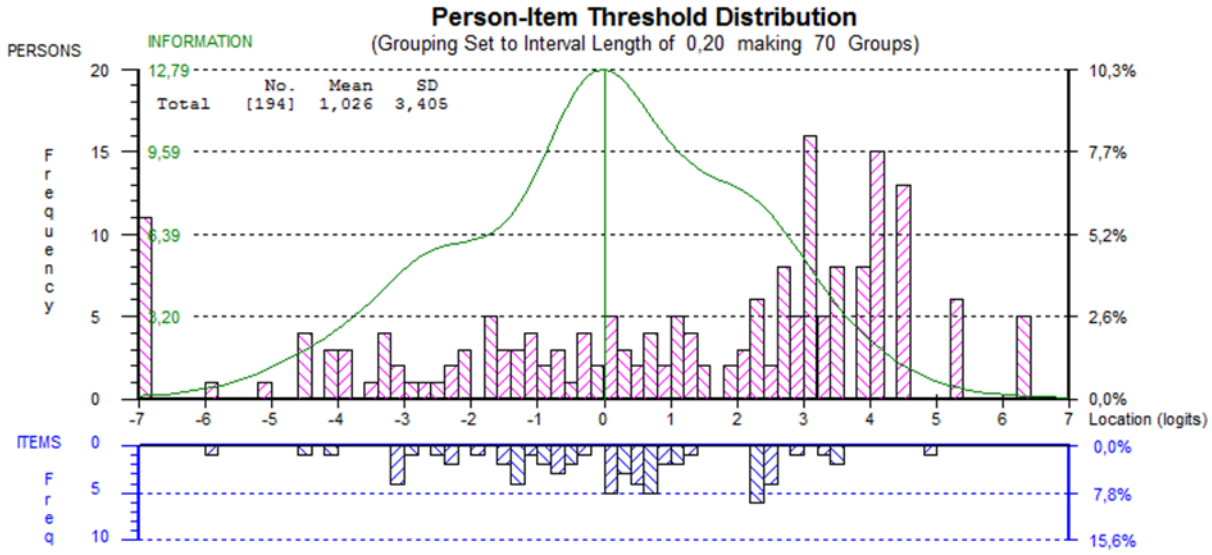


Figure 3: Threshold map for items in ranked order of difficulty according to Rasch analysis

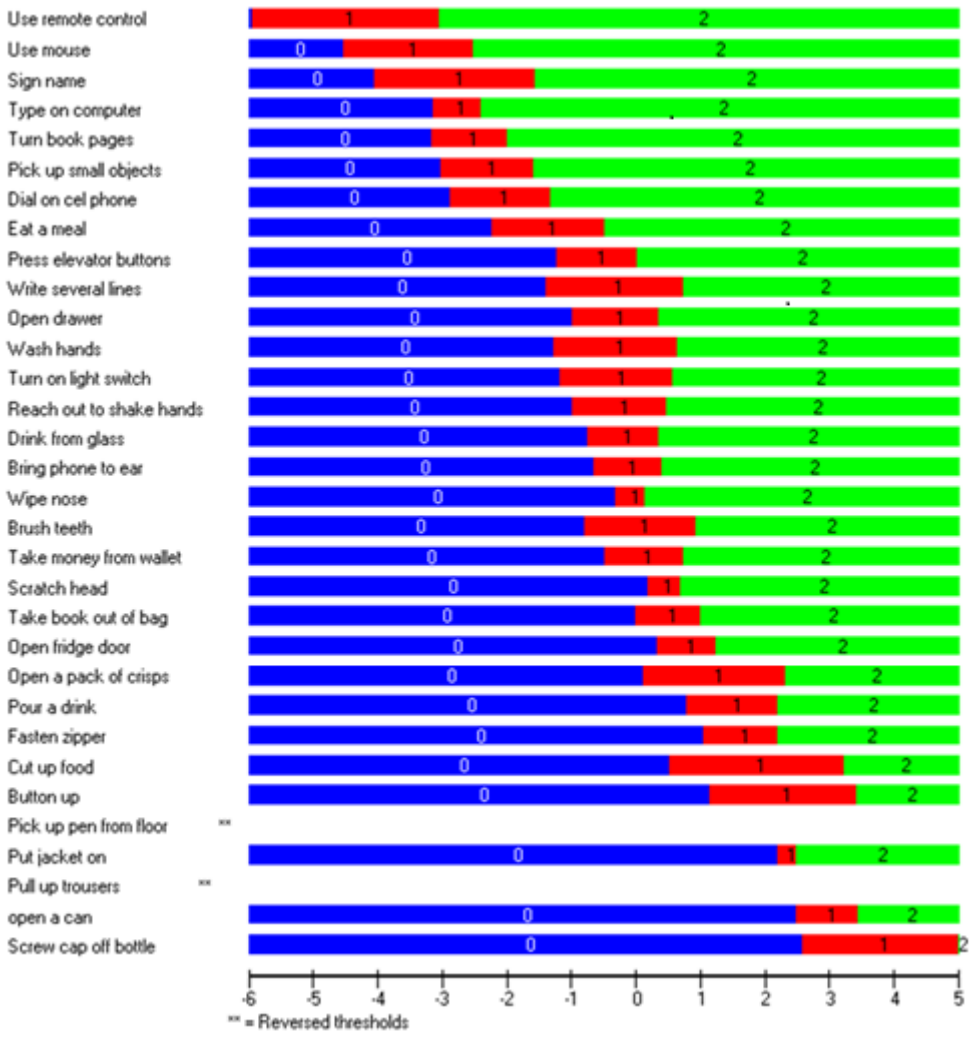
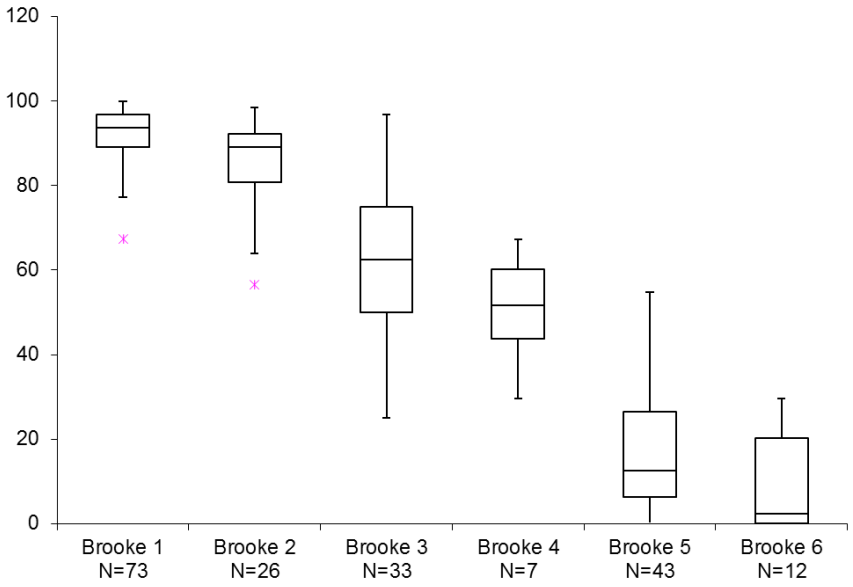


Figure 4: This box plot presents the distribution of PROM scores for the different Brooke levels. Boxes present medians and interquartile range (IQR). End of whiskers are set at 1.5*IQR above the third quartile and 1.5*IQR below the first quartile. Asterisks show the minimum of the outliers.



Legend Brooke

1 Starting with arms at the sides, can lift both arms sideways in a full circle until they touch above the head
2 Can raise both of arms above head only by flexing elbow (i.e., shortening the circumference of the movement) or using trick movements
3 Cannot raise hands above head but can raise an 8-oz. (250 ml) glass of water to mouth (by using one or both hands)
4 Can raise hands to mouth (can raise each hand separately) but cannot raise an 8-oz. (250 ml) glass of water to mouth
5 Cannot raise hand to mouth but can use hands to hold a pen or pick up coins from the table
6 Cannot raise hands to mouth and has no useful function of hands

Item	Item location	Standard error	Fit residual	χ2	χ2 *p* values
Use remote control	-4,50	0,29	-0,77	1,81	0,40
Use mouse	-3,53	0,25	0,81	7,30	0,03
Sign name	-2,80	0,23	-0,71	0,34	0,84
Type on computer	-2,78	0,23	-1,15	2,59	0,27
Turn book pages	-2,58	0,22	-1,21	1,24	0,54
Pick up small objects	-2,31	0,22	-0,76	0,47	0,79
Dial/text on phone	-2,10	0,21	1,66	58,97	0,00
Eat meal	-1,37	0,20	1,16	0,34	0,84
Press elevator buttons	-0,61	0,19	-0,72	1,91	0,39
Write lines	-0,33	0,19	2,33	36,92	<0,0001
Open a drawer	-0,32	0,19	-1,63	1,51	0,47
Wash hands	-0,31	0,18	-1,73	2,21	0,33
Turn on a light switch	-0,30	0,18	-0,52	0,70	0,71
Reach out shake hands	-0,25	0,18	-1,20	2,66	0,26
Drink from a glass	-0,19	0,18	-1,48	2,94	0,23
Bring phone to ear	-0,12	0,18	-1,86	5,29	0,07
Wipe nose	-0,08	0,18	-1,30	5,39	0,07
Brush teeth	0,07	0,18	-1,90	4,19	0,12
Take money from wallet	0,13	0,18	0,22	1,57	0,46
Scratch head	0,44	0,18	-1,40	5,17	0,08
Take book out of bag	0,50	0,18	-1,88	3,31	0,19
Open fridge	0,79	0,17	-0,85	1,10	0,58
Open crisps	1,22	0,17	0,07	1,47	0,48
Pour drink	1,50	0,16	-1,94	6,16	0,05
Fasten zipper	1,62	0,16	-1,07	1,94	0,38
Cut up food	1,87	0,16	-0,55	0,10	0,95
Button up	2,29	0,16	-0,36	1,30	0,52
Pick up pen from floor	2,31	0,15	-0,25	1,19	0,55
Put jacket on	2,35	0,15	-1,20	5,73	0,06
Pull up trousers	2,63	0,14	-0,31	2,72	0,26
Open can	2,97	0,15	-0,14	4,99	0,08
Screw off cap	3,79	0,16	-0,34	0,81	0,67

All chi-squared statistics have two degrees-of-freedom.

Table I: Individual item fit for the 32 items in threshold location: order of difficulty, easiest to most difficult.

Appendix S1: Upper Limb Clinical Outcome Group

M Eagle, John Walton Muscular Dystrophy Research Centre, Institute of Genetic Medicine, Newcastle Upon Tyne, UK;

I De Groot, University Medical Centre, Nijmegen, the Netherlands;

M Main, Dubowitz Neuromuscular Centre, UCL Institute of Child Health, London, UK;

S Messina, Department of Neurosciences, University of Messina, Italy;

G Campion, Biomarin, Leiden, the Netherlands;

L Servais, Institut de Myologie, GH Pitié Salpêtrière, Paris, France;

M Van den Hauwe, University Hospitals Leuven, Department of Child Neurology, Leuven, Belgium;

E Es, University Hospitals Leuven, Department of Child Neurology, Leuven, Belgium;

P M Pane, Department of Paediatric Neurology, Catholic University, Rome, Italy;

F Buccella, Duchenne Parent Project, Italy;

J Kuijer, Duchenne Parent Project, the Netherlands;

F Ceradini, Duchenne Parent Project, Italy;

E Henricson, Department Physical Medicine and Rehabilitation, University of California, Davis Medical Center, Sacramento, California, USA;

C McDonald, Department Physical Medicine and Rehabilitation, University of California, Davis Medical Center, Sacramento, California, USA;

J Florence, Department of Neurology, Washington University School of Medicine in Saint Louis, Missouri, St. Louis, USA