ACTIVLIM: A Rasch-built measure of activity limitations in children and adults with neuromuscular disorders

Laure Vandervelde [a], Peter Y.K. Van den Bergh [b],
Nathalie Goemans [c], Jean-Louis Thonnard [a,*]

[a] *Laboratory of Rehabilitation and Physical Medicine, Université catholique de Louvain,*
Unité de Réadapatation et de Médecine Physique (READ 5375), Tour Pasteur, 53, Av Mounier, Brussels 1200, Belgium
[b] *Neuromuscular Center and Neurology Department, Cliniques universitaires Saint-Luc, Brussels, Belgium*
[c] *Neuromuscular Center and Department of Child Neurology, Universitair Ziekenhuis Gasthuisberg, Leuven, Belgium*

Abstract

A common measure of activity limitations for both children and adults with neuromuscular disorders was developed using the Rasch model. A self-reported questionnaire containing daily activities was submitted to 245 adult patients and to the parents of 124 affected children from the two major Belgian communities. They were asked to provide their perceived difficulty in performing daily activities on a three-level scale. The 22 items of the final scale define a unidimensional and linear measure of activity limitations and show a continuous progression in their difficulty. The item difficulty hierarchy is invariant with regard to the diagnosis, community, gender and age. The scale exhibits a good precision, since the 22 items are well targeted on our sample ($r = 0.96$); furthermore, it is reproducible over time ($ICC = 0.93$). The patients' measures are related to the Functional Independence Measure motor score ($\rho = 0.85$), to the Brooke ($\rho = -0.63$) grade and to the Vignos ($\rho = -0.83$) grade.
© 2007 Elsevier B.V. All rights reserved.

1. Introduction

Most neuromuscular disorders (NMD) have a progressive clinical course that is characterized by a decrease of muscle strength [1] leading to an impaired motor function. Some consequences are fatigue, problems with locomotion and loss of functionality in activities of daily living. The International Classification of Functioning, Disability and Health (ICF) describes an individual's functioning in three domains taking into account his health condition [2]. These domains are (1) body functions and anatomical structures, (2) activity, and (3) participation. Problems in each domain are, respectively, impairments, activity limitations, and participation restrictions. In NMD patients, impairments such as muscle weakness are frequently assessed using quantitative or manual testing [3,4]. However, the evaluation of the functional abilities of patients can be also considered as a priority. These could be assessed from the level of activity limitations defined as the difficulties a patient may have in executing daily activities [2]. The achievement of daily activities depends on the muscle strength, but the relationship between the two is not straightforward [5]. It is a combination of motor function, compensatory behaviour of the patient, and personal (e.g. age, lifestyle, motivation) and

* Corresponding author. Tel.: +32 2 764 53 67; fax: +32 2 764 53 60.
E-mail address: Thonnard@read.ucl.ac.be (J.-L. Thonnard).

environmental (e.g. architectural characteristics, ground type) factors. For these reasons, the activity level should be evaluated separately and not simply inferred from the patients' impairments.

Instruments specifically applicable to the population being studied are essential for clinical evaluation [6], and a common scale for both children and adults makes it possible to follow patient status across time. The existing scales measure the functional status of NMD patients, either in a restrictive and general way, with a description of patients' limb function on a single grade (e.g., Vignos or Brooke grades [3]) or they do not measure the activity limitations themselves. The Functional Independence Measure [7] takes into account technical or human assistance and gives a measure of the patients' autonomy. A motor function measure was recently developed and validated for NMD paediatric and adult patients [8]. This scale proposes a motor measure focused on the observation of analytical tasks achieved by the patients. The time wasted for observation could be reduced by self-reported measures, especially since observed functional abilities are not psychometrically superior or easier to administer than reported measures [9]. Self-reported measures in adult patients are usually considered the gold standard [10]. In child functional assessment, parents are valid proxies since they report a finer perception of their children's abilities than the children themselves do [11,12]. The purpose of the study is to develop ACTIVLIM, a self-reported questionnaire of activity limitations in children and adults with any NMD by submitting it to the adult patients and to the parents of the affected children.

2. Patients and methods

2.1. Patients

This multicentric study was approved by the Medical Ethics Committees of the Université catholique de Louvain and of the Katholieke Universiteit Leuven. The patients were recruited through the neuromuscular reference centres of two university hospitals, each in a different Belgian language community (Dutch and French). Moreover, 10% of the children came from three centres specializing in NMD. Adult patients and parents of affected children gave written informed consent before the evaluation.

Age, gender, language community, type of NMD, Functional Independence Measure motor score [7], and Vignos and Brooke Grades [3] were included as independent demographic and clinical indices in the validation analysis. Three hundred and sixty-nine patients (124 children and 245 adults) with a diagnosis of neuromuscular disorder were assessed by the same examiner. Sample description is given in Table 1.

2.2. Questionnaire development

The ACTIVLIM questionnaire was designed to cover the widest range of daily activities and it included activities for children and for adults. The preliminary questionnaire included 138 items selected from various existing scales: ABILHAND [13,14], ABILHAND-Kids [11], Physical Functioning Subscale of SF-36 [15], Sickness Impact Profile [16], Amyotrophic Lateral Sclerosis Functional Rating Scale [17], ADL Self-Report [18], Paediatric Quality of Life Inventory [19], Lower Extremity Functional Scale [20], EK Scale [21] and Level of Rehabilitation Scale III [22]. These items were submitted to 32 experts on patients with NMD (neurologists, physicians, physical therapists, occupational therapists, nurses and a psychologist) and to 23 NMD adult patients. The experts were asked to determine the relevance of the activities for a NMD child and for a NMD adult. The adult patients were asked to evaluate the perceived difficulty in performing each activity. Both experts and adult patients were asked to propose other relevant items not included in the original item set.

The questionnaire for adult patients was achieved by removing 52 items, either because experts considered them irrelevant (44 items), or because the analysis of the 23 adults' responses through the Rasch model showed that they did not contribute to the definition of a unidimensional variable (8 items) [23]. Five items were added to the set following experts' and patients' suggestions. The adult patients were therefore assessed with a 91-item experimental questionnaire.

The experimental questionnaire for children included 99 items, since 39 items from the original set of 138 items were eliminated because experts considered them to be irrelevant.

2.3. Instrument

The ACTIVLIM questionnaire explored difficulties of performing daily activities that required the use of upper limbs or/and the use of lower limbs. The adult patients and the parents of affected children filled in either the adult form or the child form of the questionnaire. They were asked to provide their perceived difficulty in performing each activity using a three-level scale: impossible (0), difficult (1), easy (2). Each activity must be completed without technical or human assistance. Activities unfamiliar to individual patients were recorded as missing responses (2.2% of the data).

2.4. Procedure

A French or a Dutch version of the questionnaire was presented to patients. The questionnaires were self-completed by the adults or by the parents of the affected children either during their multidisciplinary

Table 1
Patient sample description ($n = 369$)

	Adults ($n = 245$)	Children ($n = 124$)
Gender		
Male (%)	56	68
Female (%)	44	32
Age, years: mean (range)	47 (16–80)	10 (6–16)
Spoken language		
Dutch (%)	42	64
French (%)	58	36
Diagnosis[a]		
DMD/BMD or LGMD (%)	15	38
HN (%)	16	28
MD (%)	17	
ALS (%)	9.5	
SMA (%)	5.5	11
FSHD (%)	5	
Others (e.g, CM, CMD, PPS, ...) (%)	32	23
Mobility level		
Walking (%)	70	66
Wheelchair-bound (%)	30	34
FIM[b]		
Motor score, median (range)	80 (25–91)	81 (31–91)
Physical therapy		
No (%)	34	11
Yes (%)	64	89

[a] DMD, Duchenne muscular dystrophy; BMD, Becker muscular dystrophy; LGMD, limb-girdle muscular dystrophy; HN, hereditary neuropathy; MD, myotonic dystrophy; ALS, amyotrophic lateral sclerosis; SMA, spinal muscular atrophy; FSHD, facio-scapulo-humeral dystrophy; CM, congenital myopathy; CMD, congenital muscular dystrophy; PPS, post-polio syndrome.
[b] FIM = Functional Independence Measure.

consultation at the Neuromuscular Centres or in the specialized centres for NMD children. The items were randomly presented to avoid an effect caused by the item order. Two hundred and twenty-seven patients were assessed a second time three weeks after the first evaluation in order to investigate the test–retest reliability of the scale.

2.5. Data analysis

To calibrate a common scale for both children and adults with NMD, responses of the parents and those of adult patients should form a single matrix. From the 91-item questionnaire for adult patients and from the 99-item questionnaire for children, 64 items were identical in both questionnaires. The remaining 27 items specifically designed for adults (91 items – 64 items) were recorded as a missing response in the data of the parents of the affected children. Similarly, the 35 specific items for children (99 items–64 items) were recorded as a missing response in the data of the adult patients. The final data matrix therefore included 126 items (64 common, 35 children specific and 27 adult specific items) that were analysed with the Rasch Unidimensional

Measurement Models computer program (RUMM2020, RUMM Laboratory Pty Ltd., Perth, Western Australia).

Since the number of adults involved in the study was twice as large as the number of children, the responses of the adult patients could have a significant effect on the scale calibration. To remedy this problem, the adult sample was divided into two equal stratified samples. Responses of the first adult sample were analysed with the responses of the parents in order to calibrate the ACTIVLIM scale (sample-1). Analysis of responses of the second adult sample with those of the parents was useful for validating the scale (sample-2).

2.6. The Rasch model

The Rasch model estimates the item difficulty and the patient activity level on a common linear scale [23,24] from the responses given to each item within a probabilistic framework [25]. This model is used to investigate response category functioning, scale unidimensionality, patient targeting and scale reliability [26]. The category functioning is studied by verifying that successive response categories for each item represent increasing levels of activity and that thresholds between adjacent categories are located in the expected order. The thresholds correspond to the activity levels required to have a higher probability to select a category rather than the previous one. The Rasch model also makes it possible to verify that all items contribute to the definition of the unidimensional activity construct [27]. To test unidimensionality, the sample is divided along the variable into level groups called class intervals. For each item, the degree of similitude between the observed responses in each class interval and the expected responses predicted by the model is computed through a standardized residual and a χ^2 fit statistic reported by the software [28]. The standardized residual of an item corresponds to the sum of the differences between the observed and the expected scores over each patient, divided by the standard deviation of the differences. It is sensitive to item discrimination. Positive values represent an under-discrimination of the item; whereas negative values represent an over-discrimination of the item. The χ^2 fit statistic represents the deviations from the model expectations. The targeting is checked by comparing the mean patient location to the mean item difficulty. The software also reports the reliability that indicates the level of measurement precision attained.

2.7. Item selection

Starting from the 126 experimental items, indices reported from successive analyses were used to select the items that constituted the final ACTIVLIM scale. If an item did not present the following criteria, it was removed from the experimental set.

1. *An ordered rating scale*. Adults and parents were asked to provide their perceived difficulty on a three-level response scale: impossible (0), difficult (1) or easy (2). The thresholds between adjacent categories should be located in an increasing order, indicating that categories were well discriminated. When the thresholds were reversed, the rating scale did not function as expected. Any item presenting disordered thresholds was deleted.
2. *The response categories have the same discrimination across all items*. To apply a rating scale model to the data and thus to make the clinical interpretation of the scores easier, each response category must be discriminated in the same way through all items [29,30]. Items presenting discrimination significantly different from the average (Z-test) were removed. Moreover, the discrimination of the categories was compared between the adults and the parents of the affected children using a *t*-test.
3. *All items fit a unidimensional scale*. Fit statistics (standardized residual and χ^2 statistic) were used to detect items that did not contribute to the definition of a unidimensional variable. Items were deleted when they presented values below −2.5 or above 2.5 for the standardized residuals [31] or when their *p*-value of the χ^2 statistic was below 0.05.
4. *No item presents a differential item functioning (DIF)*. The invariance of item difficulty hierarchy was checked with regard to four dichotomous patient-related factors: gender (male or female), patient category (child or adult), language community (Dutch or French speaker) and type of NMD (proximal or distal NMD). A DIF can be detected by a two-way analysis of variance for each item by comparing scores across each level of patient-related factor and across levels of the activity construct, i.e. the class intervals [32]. A significant main effect for the patient-related factor shows the presence of a uniform DIF; whereas a non-uniform DIF corresponds to a significant interaction effect (patient-related factor × class interval). Items that present a uniform or/and a non-uniform DIF were removed.
5. *No redundancy in item location*. Since each response category is discriminated in the same way through the items, the two thresholds of the items are equidistant. The thresholds of two items with the same difficulty will therefore have the same location along the variable without increasing the scale reliability. When items were redundant, one of them was deleted, preferably keeping items common to adults and children.

2.8. Scale validity

A Rasch analysis with sample-2 responses on the selected activities made it possible to validate item difficulty hierarchy and scale psychometric properties. A comparison of item hierarchy between sample-1 and sample-2 was carried out using a DIF test [33]. Second, the construct validity was tested by examining the degree of association between the ACTIVLIM measures of patients and the widely used scales [34]: Functional Independence Measure motor score, Vignos and Brooke grades. Moreover, relationships between measures of patients and demographic (age, gender, community) and clinical (type of NMD) indices were studied to examine consistency with plausible hypotheses [35]. A correlation coefficient was computed for continuous indices and a one-way analysis of variance for groups of nominal indices.

2.9. Scale reliability

A reliability index was determined as the proportion of observed measure variance attributable to the true measure variance [36]. Moreover, the test–retest reliability of the adults' and the parents' responses was determined by the intraclass correlation coefficient [37]. A DIF test was carried out to verify the invariance of item difficulty hierarchy across the first and the second assessment.

3. Results

3.1. Refinements of the ACTIVLIM scale for NMD children and adults

The successive analyses through the Rasch model selected 22 items from the original 126-item set. Seven items presented reversed thresholds, 14 items had a different discrimination of the response categories, 49 items did not contribute to the definition of the unidimensional construct, 22 items showed a uniform or a non-uniform DIF (2 items with regard to the gender patient-related factor, 9 for the patient category, 4 for the language community and 7 items for the type of NMD) and 12 items were redundant. Moreover, the discrimination of the categories was not significantly different in adults and children's parents ($t = 0.179$, $p = 0.86$). The final ACTIVLIM scale contains 14 common activities, 4 specific activities for children, and 4 specific activities for adults (Table 2).

3.2. Psychometric properties of the ACTIVLIM scale in NMD children and adults

The calibration of the final 22-item ACTIVLIM scale is presented in Table 2. The items are classed in decreasing difficulty order from top to bottom (range: 3.57 to −3.33 logits), with higher logit values representing more difficult items. The activity measure is expressed in logits (i.e. log-odds units), a linear unit

Table 2
ACTIVLIM calibration for adults and children with neuromuscular diseases

Items[a]			Difficulty (logits)	SE (logits)	Residual (z)	Fit (χ^2)	Degrees of freedom[b]	p-value
a	Hopping on one foot	C	3.57	0.23	0.06	6.20	4	0.18
b	Carrying a heavy load	A	3.03	0.20	−0.43	1.79	4	0.77
c	Running	C	2.78	0.21	−0.56	0.52	4	0.97
d	Walking more than 1 km	A	2.29	0.18	−0.61	4.17	4	0.38
e	Walking upstairs	B	1.47	0.13	−1.86	10.11	9	0.34
f	Standing for a long time	A	1.30	0.18	0.41	0.72	4	0.95
g	Stepping out of a bath tub	B	1.18	0.14	−0.61	9.01	9	0.44
h	Walking downstairs	B	1.03	0.13	−1.50	3.07	9	0.96
i	Taking a bath	B	0.27	0.14	−1.02	10.39	9	0.32
j	Putting on a backpack	C	0.04	0.21	−0.92	5.12	4	0.28
k	Dressing one's lower body	B	−0.23	0.14	−1.19	7.74	9	0.56
l	Walking outdoors on level ground	B	−0.26	0.14	−1.10	10.45	9	0.32
m	Getting into a car	A	−0.33	0.19	−1.13	5.23	4	0.26
n	Taking a shower	B	−0.45	0.15	−1.10	10.84	9	0.29
o	Wiping one's upper body	B	−0.86	0.15	0.77	8.53	9	0.48
p	Putting on a T-shirt	B	−1.19	0.16	−0.12	7.10	9	0.63
q	Hanging a jacket on a hat stand	B	−1.28	0.16	−0.48	3.62	9	0.93
r	Sitting on the toilet	B	−1.49	0.16	−1.01	10.97	9	0.28
s	Washing one's upper body	B	−1.64	0.17	−0.45	14.59	9	0.10
t	Opening a door	B	−2.73	0.19	0.37	14.22	9	0.11
u	Closing a door	C	−3.16	0.26	−0.54	2.38	4	0.67
v	Washing one's face	B	−3.33	0.21	0.54	9.57	9	0.39

[a] A, B or C corresponds respectively to an adult, to a common or to a child activity.
[b] Degrees of freedom are the number of class interval (CI)−1; the specific items were composed of 5 CI while the common items were composed of 10 CI in order to balance the number of subjects per CI (±23 subjects).

defined as the natural logarithm of the odds of successful achievement by a patient for any item. This unit is constant along the measurement scale and the zero of the scale is set by convention at the average difficulty of the whole selected item set. The table also shows the standard error (SE) of the item difficulty estimates (mean: 0.17, range: 0.13–0.26), the standard residual (mean: −0.57, range: −1.86–0.54), the fit statistic computed as a χ^2 and the associated p-value. All 22 items define a unidimensional scale of activity limitations in NMD patients since p-values do not show a significant difference between observed and expected scores. The reliability index of the final scale is equal to 0.96, indicating that 7 groups of activity levels can be statistically distinguished within the patient sample [38].

3.3. Item and psychometric properties validity

The Rasch analysis of the 22 final items on sample-2 responses gives identical psychometric properties: item difficulty range from 3.57 to −3.5 logits, mean SE of item difficulty estimates is equal to 0.17, all the items present a non-significant p-value for fit statistics, and the reliability index is equal to 0.96. Moreover, the differential item functioning plot presented in Fig. 1 compares the item difficulty hierarchy as estimated by the sample-1 and the sample-2. The 22 items lie within the 95% of confidence interval of the identity line, indicating that the items were consistently estimated by both samples.

3.4. Description of the ACTIVLIM scale

Fig. 2 shows the structure and the targeting of the final linear ACTIVLIM scale in NMD patients. The overall mean patient location is equal to 0.7 logits, indicating that the selected items are well targeted to the NMD sample. Moreover, the range of measurement represented by the thresholds distribution fits the distribution of the patients' abilities. According to the distribution of patients' measures, adult patients with a measure above 1.83 logits and children with a measure above 2.36 logits should achieve all activities easily or with some difficulty.

3.5. Construct validity

No significant differences in ACTIVLIM measures were observed with regard to age ($r = -0.04$, $p = 0.428$), language community ($F = 0.294$, $p = 0.588$), gender ($F = 0.004$, $p = 0.95$), and patient category ($F = 0.31$, $p = 0.654$). A difference in ACTIVLIM measures was found with regard to the type of NMD ($F = 15.92$, $p < 0.001$). A post-hoc analysis indicates that the patients with proximal NMD have a lower activity level than do patients with other types of NMD. The ACTIVLIM measures were correlated with the Functional Independence Measure motor score ($\rho = 0.85$), the Vignos ($\rho = -0.83$), and the Brooke ($\rho = -0.63$) grades (Fig. 3).
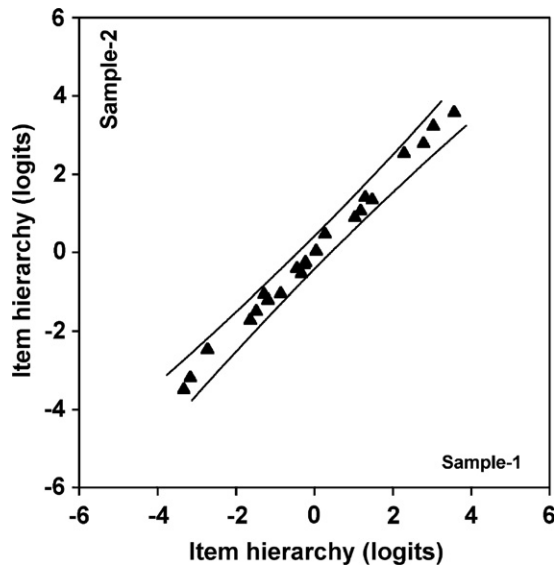
Fig. 1. Item difficulty hierarchy estimated by sample-1 and sample-2. Most difficult items are plotted in the upper right part of the figure. Control lines (solid lines) indicate the 95% confidence intervals of the ideal invariance. All items (dots) lie within the control line, indicating that both samples estimated the item difficulty consistently.

3.6. Test–retest reliability

The test–retest reliability (delay: 24 ± 9 days) of the patient measures is shown in Fig. 4 (right panel). Most of the measures lie within the 95% CI of the identity line, indicating that adults and parents tend to consistently estimate their own or their child's activity level. Moreover, the ICC for the patient measures is equal to 0.93. The left panel shows the DIF plot of the item difficulty hierarchy between the first and the second assessment. Two items identified by their labels lie outside the 95% CI of the identity line. The ICC of the item estimates is equal to 0.98, indicating a good reproducibility of the item hierarchy.

4. Discussion

The purpose of this study was to develop a common measure of activity limitations using the Rasch model and to validate it in both adults and children with NMD. The ACTIVLIM questionnaire was constructed from the adults' and the parents' perception of the difficulty in performing activities of daily living. The 22 items selected for the final version of the ACTIVLIM scale share the same ordered rating scale structure, fit a unidimensional scale and present no differential item functioning across age, gender, speech community and type of NMD.

The few items removed because of reversed thresholds indicate that both adults and the parents of affected children correctly discriminate the three proposed response categories [39]. Moreover, a common rating

scale model for children and adults could be applied since the response categories were equally discriminated by the adults and the parents of the affected children [29]. The perception of the parents was indeed preferred to that of the children's because children have a more dichotomous perception of their abilities [12]. They perceived the activities either as "impossible" or "easy" with rare intermediate responses [11]. Therefore, the use of the children's responses could lead to a narrower range of measurement, with more patients with extreme measures, leading to a less appropriate scale for the NMD sample.

Despite the temptation to construct a scale measuring different facets of NMD patient (e.g. activity limitations, fine hand motor skills, cognition,...), ACTIVLIM is a unidimensional scale that only measures activity limitations without other characteristics potentially leading to biased results of the evaluation [23]. The large number of items that did not contribute to the definition of the unidimensional variable may indicate that the experimental questionnaire measures more than one variable [27]. Seventy percent of such deleted items mainly require hand and finger strength in order to be achieved (e.g., cutting meat, fastening the snaps of a jacket, unscrewing a bottle cap) and they appear to assess the manual ability of the patients. The reasons for the unsuitability of these items to the model can be explained after data examination [40]. The patients with a distal NMD have more difficulties in performing the manual activities than do patients with a proximal NMD; whereas the former have a higher level of activity. For this reason, the scores observed for the manual items do not correspond to the scores predicted by the model. The final ACTIVLIM scale does not include exclusively digital and manual activities, but it is suitable for all types of NMD and is reliable enough to be clinically useful. A scale of manual ability is however being developed for NMD patients.

The differential item functioning tests allowed to select items with no significantly different hierarchy between the compared person-related factors. For example, the hierarchy of the 22 selected items is invariant if the item difficulty is estimated by patients with a proximal NMD or by patients with a distal NMD. The same invariance was also observed between males and females and between Dutch and French speakers for the 22 final items, and between adults and children for the 14 common items.

The hierarchy of the 22 items retained for the scale is consistent with the psychomotor qualities and with the energy expenditure necessary to perform the activities. The activities requiring more balance, force or endurance, and therefore higher energy expenditure [41], tend to be more difficult for NMD patients. The easiest activities can be often managed in a sitting position, using adaptive strategies. The most difficult
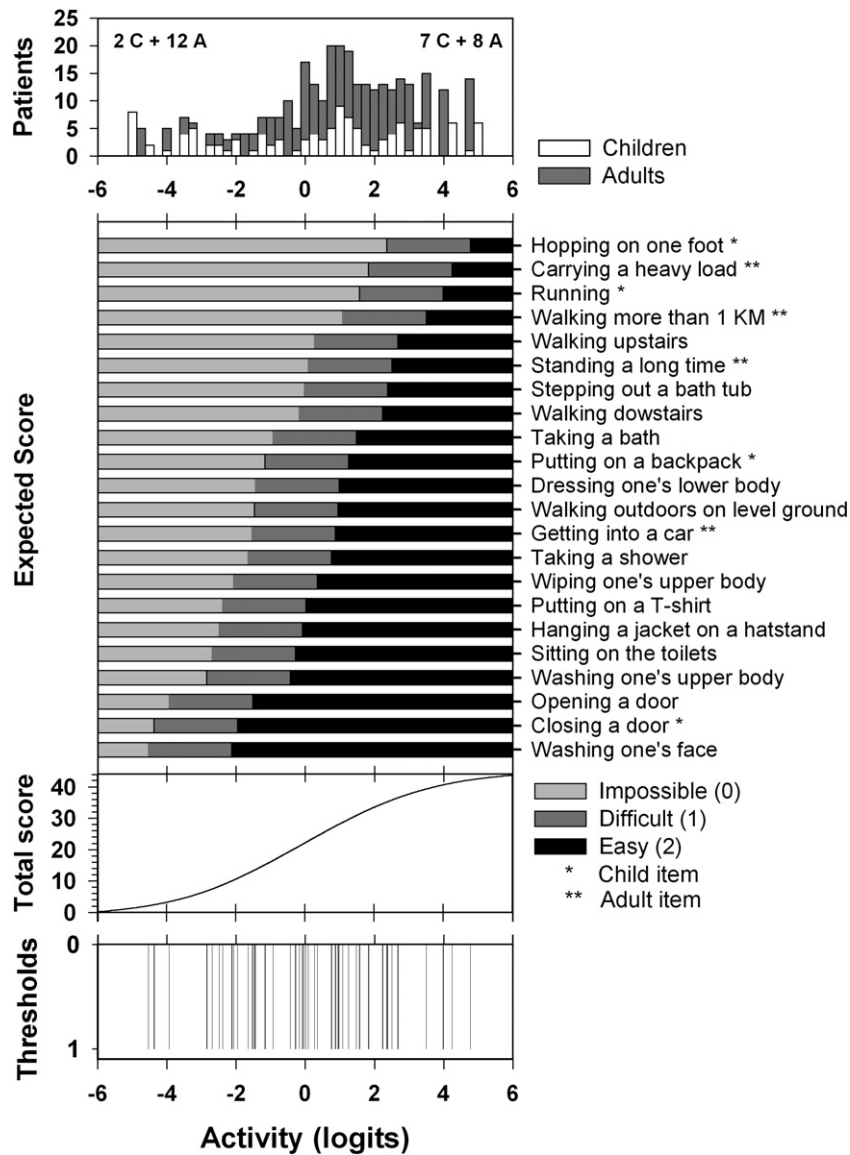
Fig. 2. Top panel: Distribution of activity measures of the 369 NMD patients, the children's measures as perceived by their parents are in white and the self-perceived adults' measures are in grey. Nine children (C) and 20 adults (A) with extreme scores cannot be assessed by the activity scale because all activities were either impossible (2 C + 12 A) or easy (7 C + 8 A). Upper middle panel: A patient's expected response for each item as a function of the activity measure. Zero is, by convention, set at the average item difficulty. For example, being able to easily put on a T-shirt requires a measure of at least 0.03 logits; whereas any patient with a measure below −2.39 would be unable to fully accomplish this activity. Conversely, a patient with a measure of 0 logits would be expected to perform the six easiest activities without difficulty, to perform the average activities with some difficulties, and would be unable to perform the six most difficult activities. Lower middle panel: Ogival relationship that makes it possible to convert the ordinal total score into a linear activity measure. Total score ranges from 0 to 44 because 22 items scored from 0 to 2. The activity measure in logits is an infinite measure. Although this relationship is quasi-linear in its central part, a unitary increment in total score encompasses larger differences of activity measure at both ends of the scale. Bottom panel: Scale graduations represented by the 44 thresholds (2 thresholds by item). The first graduation corresponds to the activity level required to endorse "difficult" rather than "impossible" for the item "washing one's face" and the last one represents the activity level required to select "easy" rather than "difficult" for the item "hopping on one foot".

activities usually involve lower limbs, and wheelchair-bound patients would answer "impossible" for these activities. As "impossible" correspond to "0", total raw score of wheel-chair bound patients will be lower than the one of walking patients, as well as their activity level expressed in logits (Fig. 2, third panel). In addition, the hierarchy of the 22 selected items was consistently estimated by both samples of patients (Fig. 1) and the

psychometric properties of the sample-2 scale are equivalent to the calibration of sample-1. This indicates that the selected items correspond to pertinent and appropriate activities to assess activity limitations in any NMD patients.

Among the 22 final items, 4 items specifically evaluate children with NMD and 4 specifically evaluate adults with NMD. Following experts' advices, the 4 specific
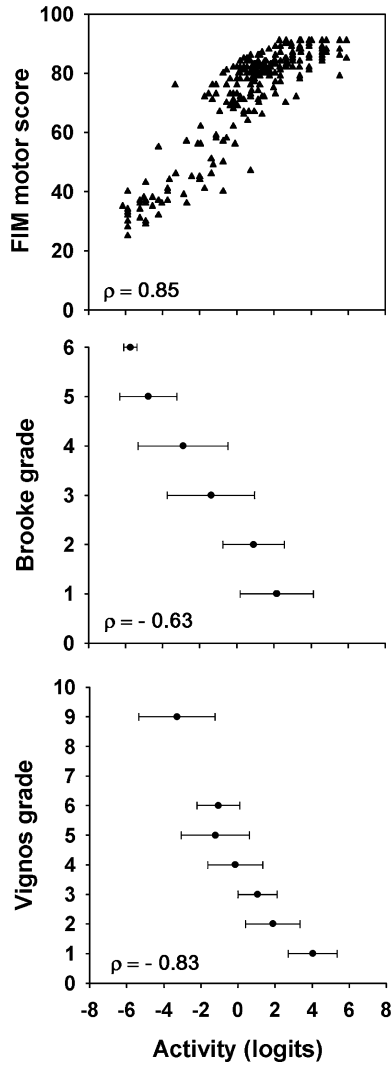
Fig. 3. Relationship between patients' activity measures and the motor score of the Functional Independence Measure (top panel), the Brooke (middle panel), and the Vignos grades (bottom panel).

items for adults (items b, d, f and m) were not relevant for a child evaluation while 3 specific items for children (item a, c, and j) were not relevant for an adult evaluation. The fourth child item (item u) did not fit the unidimensional construct of activity limitations in the first Rasch analysis based upon the responses of 23 adult patients.

The standard errors associated with the item difficulty estimates (mean: 0.17 logits) conform to the expectation for most variables and are low enough to make the measurement precision high [42]. The ACTIVLIM scale presents good reliability since the 44 graduations are well targeted on our sample, representing a wide range of functional states ($R = 0.96$). Only 4% of the patients were unable to perform at least one activity. Most of these patients had a proximal NMD and all of them were wheelchair-bound (Vignos grade of 9) with an extremely affected upper limb function (Brooke grades of 5 or 6). Likewise, 4% of the patients were able to easily complete all the activities. Most of these patients had a distal NMD or myotonic dystrophy. All of them were able to raise their arms above their heads without flexing the elbows (Brooke grade of 1) and were able to walk and to go upstairs without assistance or using a railing (Vignos grades of 1 or 2). This low percentage of patients located at the ends of the scale indicates that the scale has no significant ceiling or floor effect [43]. The wide range of scale graduations and the 22 items are sufficient to measure activity limitations in patients with any diagnosis of NMD.

The difficulty of two items ("closing a door" and "putting on a T-shirt") slightly differs between the first and the second assessment (Fig. 4). However, the differential item functioning of the two items is not high enough to compromise the test–retest reliability of the questionnaire. The high intraclass correlation coefficient
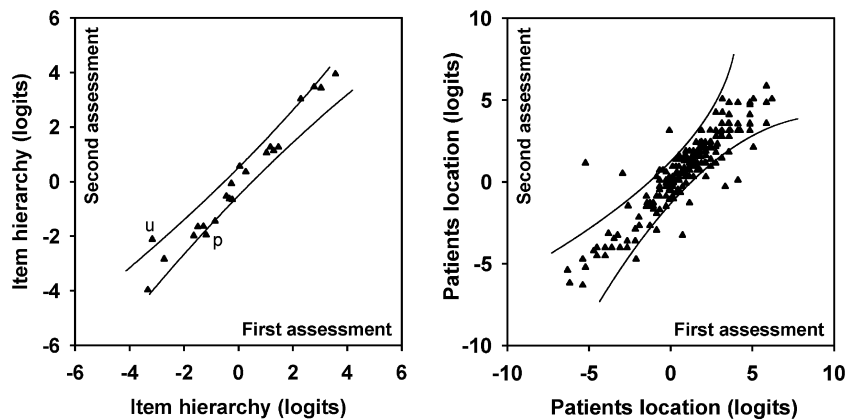

Fig. 4. Left panel: Differential item functioning plot of the item difficulty perceived by the adults and the children's parents at the first and the second assessment (delay: 24 ± 9 days) and the 95% of confidence interval (solid line) of the ideal invariance. Most difficult items are plotted in the top/right part of the figure. Two items lying outside the CI are identified by their label. Right panel: Relationship between the activity measure of patients assessed across time and the 95% of CI (solid lines) of the ideal invariance. More active patients are plotted in the top/right part of the figure. Patients' measures (dots) lying within the control lines have the same estimated activity at the first and the second assessment.

found for the item hierarchy (ICC = 0.98) after a delay of 1 month indicates that the questionnaire is reproducible over time. Moreover, the adult patients and the parents of the affected children respectively evaluate themselves and their children consistently after about 1 month.

The analyses of the relationships between the patients' measures with other widely used scales, such as the Vignos and Brooke grades and the motor score of the FIM, highlighted the good construct validity of the scale, with correlation coefficients of −0.83, −0.63 and 0.85, respectively. A higher activity level relates to lower Vignos and Brooke grades and to a higher level of independence. The Vignos and the Brooke grades respectively class the function of the lower and the upper limbs into a single category. Nevertheless, each category represents a rather wide range of ACTIVLIM measures in logits (Fig. 3). ACTIVLIM is therefore more complete and precise than both of these grades since it allows to differentiate groups of patients within a same category of Vignos and Brooke grades. Concerning the FIM, few studies have validated it in a NMD population [44,45]; although it is one of the most commonly used questionnaires in the evaluation of NMD patients [4]. Moreover, the motor score of the FIM seems not to be precise enough to distinguish groups of patients in the high levels of the motor score (Fig. 3). Indeed, half of patients have a motor score above 80, indicating a high level of independence; while their activity measures range from −0.55 to 5.9 logits, representing a wide range of activity levels. These results confirm the observations in patients with poliomyelitis sequelae [46], among which a large number were independent in the activities of daily living, even if they reported difficulties in these activities. However, the FIM motor score and ACTIV-LIM measure different aspects of the patient's health condition. ACTIVLIM evaluates activity limitations in terms of difficulties in performing daily activities without technical or human assistance, and the FIM measures the independence of the patient taking into account the environmental factors [2]. For example, patients who can propel their wheelchair themselves are considered independent for the locomotion item "walk/wheelchair"; yet following the ACTIVLIM questionnaire, it is "impossible" for them to walk more than one kilometre (item d), since a wheelchair is considered to be a form of technical assistance. ACTIVLIM could however determine the technical assistance necessary to achieve some items. Furthermore, this scale is more precise and detailed than the FIM and it can remedy to the lack of sensitivity of the FIM.

The relationships between the patients' measures and demographic and clinical indices appear as clinical information. The significant relationship between the patients' measures and the type of NMD confirms previous reports [46,8] that patients with distal NMD and with myotonic dystrophy are less disabled in their functional status than are patients with proximal NMD and in particular those with Duchenne muscular dystrophy. The measures of the patients were not related to age, gender or speech community.

The Rasch model was used to construct and validate the ACTIVLIM scale. This particular methodology provided a measurement scale with fundamental psychometric qualities known as linearity and unidimensionality. This questionnaire has also good reliability, precision, construct validity and reproducibility. Moreover, ACTIVLIM can be used for evaluation of both adults and children with NMD making possible to follow the disease course from childhood to adulthood using a single scale. The hierarchy of the items is invariant across age, gender, language community or type of NMD indicating that ACTIVLIM can be used for any patients with NMD. Finally, the questionnaire is extremely easy to administer, since it can be completed in 5 minutes in the waiting room by the patient himself or a child's parent. Nevertheless, ACTIVLIM does not claim to replace clinical evaluation methods that principally measure the impairments (manual muscle testing, range of motion, timed tasks test, etc.) [2]; it is rather complementary to these.

Further applications of the ACTIVLIM scale include the study of its responsiveness. The high precision of the scale ensures to statistically predict a good sensitivity to change in activity limitations induced, for instance, by the progressive course of the disease or by treatment. Nevertheless it must be clinically verified. Moreover, its relationships with impairment measures and its similarity between the self-reported version and the achievement of daily activities observed by a therapist will also be investigated.

For clinical use of ACTIVLIM, a website (www.rehab-scales.org) including downloadable scoring sheets and instructions will be available soon. Moreover, on-line analyses taking into account the missing values would directly convert raw scores into linear measures of activity limitations.

References

[1] Piccininni M, Falsini C, Pizzy A. Quality of life in hereditary neuromuscular diseases. Acta Neurol Scand 2004;109(2):113–9.

[2] World Health Organisation. The International Classification of Functioning, Disability and Health-ICF. Geneva: WHO, 2001.

[3] Brooke MH, Griggs RC, Mendell JR, Fenichel GM, Shumate JB, Pellegrino RJ. Clinical trial in Duchenne dystrophy. I. The design of the protocol. Muscle Nerve 1981;4(3):186–97.

[4] Haigh R, Tennant A, Biering-Sørensen F, et al. The use of outcome measures in physical medicine and rehabilitation within Europe. J Rehabil Med 2001;33:273–8.

[5] Merkies IS, Schmitz PI, van der Meche FG, Samijn JP, van Doorn PA. Inflammatory neuropathy cause and treatment (INCAT) group. Connecting impairment, disability, and handicap in immune mediated polyneuropathies. J Neurol Neurosurg Psychiatry 2003;74:99–104.

[6] Wade DT. Measurement in neurological rehabilitation. New York: Oxford University Press; 1992.

[7] Linacre JM, Heinemann AW, Wright BD, Granger CV, Hamilton BB. The structure and stability of the Functional Independence Measure. Arch Phys Med Rehabil 1994;75(2):127–32.

[8] Bérard C, Payan C, Hodgkinson I, Fermanian J. The MFM collaborative study group. A motor function measure scale for neuromuscular diseases. Construction and validation study. Neuromusc Disord 2005;15:463–70.

[9] Myers AM, Holliday PJ, Harvey KA, Hutchinson KS. Functional performance measures: are they superior to self-assessments? J Gerontol 1993;5:M196–206.

[10] Sprangers MA, Aaronson NK. The role of health care providers and significant others in evaluating the quality of life of patients with chronic disease: a review. J Clin Epidemiol 1992;45(7):743–60.

[11] Arnould C, Penta M, Renders A, Thonnard JL. ABILHAND-kids. A measure of manual ability in children with cerebral palsy. Neurology 2004;63:1045–52.

[12] Chambers CT, Johnston C. Developmental differences in children's use of rating scales. J Pediatr Psychol 2002;27:27–36.

[13] Penta M, Tesio L, Arnould C, Zancan A, Thonnard JL. The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. Stroke 2001;32(7):1627–34.

[14] Penta M, Thonnard JL, Tesio L. ABILHAND: a Rasch-built measure of manual ability. Arch Phys Med Rehabil 1998;79:1038–42.

[15] Ware Jr JE, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. Med Care 1992;30(6):473–83.

[16] Bergner M, Bobbitt RA, Carter WB, Gilson BS. The sickness impact profile: development and final revision of a health status measure. Med Care 1981;19(8):787–805.

[17] The amyotrophic lateral sclerosis functional rating scale. Assessment of activities of daily living in patients with amyotrophic lateral sclerosis. The ALS CNTF treatment study (ACTS) phase I–II Study Group. Arch Neurol 199;53(2): 141–7.

[18] Ahlstrom G, Gunnarsson LG. Disability and quality of life in individuals with muscular dystrophy. Scand J Rehabil Med 1996;28(3):147–57.

[19] Varni JW, Seid M, Rode CA. The PedsQL: measurement model for the pediatric quality of life inventory. Med Care 1999;37(2):126–39.

[20] Binkley JM, Stratford PW, Lott SA, Riddle DL. The lower extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. North American Orthopaedic Rehabilitation Research Network. Phys Ther 1999;79(4):371–83.

[21] 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;6(3):119–34.

[22] Velozo CA, Magalhaes LC, Pan AW, Leiter P. Functional scale discrimination at admission and discharge: Rasch analysis of the Level of Rehabilitation Scale-III. Arch Phys Med Rehabil 1995;76(8):705–12.

[23] Wright BD, Linacre JM. Observations are always ordinal; measurements, however, must be interval. Arch Phys Med Rehabil 1989;70:857–60.

[24] Merbitz C, Morris J, Grip JC. Ordinal scales and foundations of misinference. Arch Phys Med Rehabil 1989;70:308–12.

[25] Rasch G. Probabilistic models for some intelligence and attainment tests. Chicago: MESA Press; 1980. Originally published in 1960 by the Danish Institute of Educational Research.

[26] Smith Jr EV, Conrad KM, Chang K, Piazza J. An introduction to Rasch measurement for scale development and person assessment. J Nurs Meas 2002;10(3):189–206.

[27] Bond TG, Fox CM. Applying the Rasch model: fundamental measurement in the human sciences. Mahwah: Lawrence Erlbaum Associates; 2001.

[28] Wright BD, Masters GN. Rating scale analysis. Chicago: MESE Press; 1982.

[29] Wright BD. Model selection: rating scale or partial credit? Rasch Measur Trans 1999;12:641–2.

[30] Linacre JM. Comparing "partial credit" and 'rating scale' models. Rasch Measur Trans 2000;14:768.

[31] Andrich D, Sheridan BE. Interpreting RUMM2020: part II, polytomous data. 3rd ed. Perth: RUMM Laboratory Pty Ltd.; 2005.

[32] 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;46:1–18.

[33] Wright BD, Stone MH. Best test design. Chicago: MESA Press; 1979.

[34] Merkies ISJ, Schmitz PIM, van der Meché FGA, Samijn JPA, van Doorn PA. Clinimetric evaluation of a new overall disability scale in immune mediated polyneuropathies. J Neurol Neurosurg Psychiatry 2002;72:596–601.

[35] Engelberg R, Martin DP, Agel J, Obremsky W, Coronado G, Swiontkowski MF. Musculoskeletal Function Assessment instrument: criterion and construct validity. J Orthop Res 1996;14(2):182–92.

[36] Traub RE, Rowley GL. Understanding reliability. Educational Measurement Issues and Practice, National Council on Measurement in Education 1991;10(1):37–45.

[37] Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. Psychol Bull 1979;86:420–8.

[38] Fisher Jr W. Reliability statistics. Rasch Measur Trans 1992;6:238.

[39] Andrich D. Category ordering and their utility. Rasch Measur Trans 1996;9:464–5.

[40] Andrich D. E444 Advanced Social and Educational Measurement. Perth: Murdoch University; 1999.

[41] Ainsworth BE, Haskell WL, Whitt MC, et al. Compendium of physical activities: an update of activity codes and MET intensities. Med Sci Sports Exerc 2000;32(Suppl. 9): S498–504.

[42] Linacre JM. Sample size and item calibration stability. Rasch Measur Trans 1994;7:328.

[43] McHorney CA, Ware Jr JE, Lu JF, Sherbourne CD. The MOS 36-item Short-Form Health Survey (SF-36): III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. Med Care 1994;32(1):40–66.

[44] Jensen MP, Abresch RT, Carter GT. The reliability and validity of a self-report version of the FIM instrument in persons with neuromuscular disease and chronic pain. Arch Phys Med Rehabil 2005;86(1):116–22.

[45] Uchikawa K, Liu M, Hanayama K, Tsuji T, Fujiwara T, Chino N. Functional status and muscle strength in people with Duchenne muscular dystrophy living in the community. J Rehabil Med 2004;36(3):124–9.

[46] Thoren-Jonsson AL, Grimby G. Ability and perceived difficulty in daily activities in people with poliomyelitis sequelae. J Rehabil Med 2001;33(1):4–11.