Introduction: Duchenne muscular dystrophy (DMD) and Spinal muscular atrophy (SMA) causes significant disability and progressive functional impairment. Readily available instruments that assess functionality, especially in advanced stages of the disease, are required to monitor the progress of the disease and the impact of therapeutic interventions. Objective: To describe the development of a scale to evaluate upper limb function (UL) in patients with DMD and SMA, and describe its validation process, which includes self-training for evaluators. Patients and Method: The development of the scale included a review of published scales, an exploratory application of a pilot scale in healthy children and those with DMD, self-training of evaluators in applying the scale using a handbook and video tutorial, and assessment of a group of children with DMD and SMA using the final scale. Reliability was assessed using Cronbach and Kendall concordance and with intra and inter-rater test-retest, and validity with concordance and factorial analysis. Results: A high level of reliability was observed, with high internal consistency (Cronbach α = 0.97), and inter-rater (Kendall W = 0.96) and intra-rater concordance (r = 0.97 to 0.99). The validity was demonstrated by the absence of significant differences between results by different evaluators with an expert evaluator (F = 0.023, P > .5), and by the factor analysis that showed that four factors account for 85.44% of total variance. Conclusions: This scale is a reliable and valid tool for assessing UL functionality in children with DMD and SMA. It is also easily implementable due to the possibility of self-training and the use of simple and inexpensive materials.
Introduction

Neuromuscular diseases (NMD) are those that affect the peripheral nervous system control resulting in muscular control loss. In pediatric age, most NMDs lack curative treatment and involve a significant functional compromise, leading to progressive disability. Any neurological rehabilitation program that helps these patients requires instruments to monitor functionality, prevent secondary disorders, provide a common language among professionals involved, and especially to evaluate the effect of different therapies.

Among pediatric NMDs, Duchenne muscular dystrophy (DMD) and spinal muscular atrophy (SMA) are the most common, with DMD as the most prevalent and both progressive and highly disabling.12-14. These diseases affect the evolution of the disease, in addition to deliver adequate final criteria that is crucial when evaluating the effect of therapeutic interventions, especially in post-loss stages, is crucial for a proper management of these patients. These instruments are of key importance and should be accessible to professionals in charge of patients with NMDs, without resulting in excessive medical expenses. These instruments should be designed with inexpensive and easily acquired materials, have manuals and instructions that promote self-learning and be available in the language of the evaluator and the population to be evaluated.

The objectives of this study are to report the development of a scale specifically made to evaluate UL function, its application in patients with DMD and SMA, and describe its validation process, which includes self-training for evaluators.

Patients and Method

Design and development of the scale

Initially, other scales specifically designed or adapted to evaluate NMDs functionality were reviewed.5,6,21-23,25-28. Afterwards, a group of experts (pediatric neurologist specialized in NMDs, pediatric neurologist specialist in neurorehabilitation, occupational therapist, kinesiologist), with extensive experience in the management of pediatric patients with NMDs, reviewed the items intended to evaluate UL functionality, choosing an initial list of 17 items. Then, 4 additional items were added, aiming to reflect limitations in activities of daily living (ADL) in non-walking patients with NMDs. The list included a total of 21 items, which were applied to 8 healthy children (between 5 and 12 years of age) and then to 4 children with NMD (between 10 and 16 years of age). Eight items were modified according to the evaluations made creating the final list.

The final version of the scale included 21 items, which were grouped into 4 dimensions, similar to the one proposed by Mayhew29. Each item has a score from 0 to 5, except 5 of them that score between 0 and 4. The scale has a total score that ranges between 0 and 120. Subsequently, a detailed manual about the application of the scale and the implementation of the necessary kit was written. In order to have a model, the scale application to a healthy adult was filmed. This
video was watched by 4 therapists, and based on their comments, modifications to the manual were made to obtain consistency between the images and the instructions.

The occupational therapist (OT), who participated in the scale design, evaluated a total of 10 patients with NMD (between 10 and 19 years of age), in 2 sessions, each separated by 2 weeks. These sessions were filmed so that the implementation of the scale could be scored by other evaluators. Prior to the scale implementation, the OT applied Barthel's index (IB) to each of the patients in the first session 30. IB index is a 10-item scale that measures functional independence in the domains of personal care and mobility. The total score ranges between 0, total dependence, to 100, total independence. The validity and reliability of IB has been clearly established 31,32.

Evaluator training

After completing the above, 5 experienced therapists working with children with neurological disabilities (2 OT and 3 kinesiologists), became self-trained in the application of the scale using the manual and video (both available at: http://www.cedeti.cl/recursos-tecnologicos/escala-de-funcionalidad/funcionalidad-enfermedades-neuromusculares/). The frequency and time for review of the material was determined by each therapist; also, they were able to ask the OT questions.

Evaluator reliability

After the self-training phase, the evaluators received the assessment of each of the children performed by the OT, in addition to a set of the scale application guidelines. Evaluators performed the assessments consecutively, finalizing one process before starting the next. Each of the 5 evaluators applied the scale to the 10 patients, for a total of 50 evaluations (5 evaluations for each patient). After 8 weeks of the first round, evaluators repeated the assessments in the same way, completing 2 evaluations (test, re-test) for each of the 10 patients.

Patients

The sample was non-randomized and formed by 10 subjects, 8 DMD and 2 SMA, who had been followed up for at least 4 years in the Neuromuscular Diseases Unit of Catholic University of Chile, all with confirmed diagnosis by genetic-molecular study. The mean age of subjects was 12.8 years (range: 9.4 to 19.1). One of the patients with SMA was female. Three of the 10 participants presented independent ambulation, DMD and under corticoid treatment. The other 5 patients with DMD had suspended corticosteroids since the inability of walking, at least a year earlier.

All participants completed their baseline assessments safely and without difficulty. The average time of the evaluation was 20 min (range: 15 to 23) and there was no evidence of fatigue in any of the subjects.

The study was approved by the ethics committee of the School of Medicine of the Pontifical Catholic University of Chile.

Statistical Analysis

Validity and reliability of the scale were evaluated with various analyzes detailed in the Results section. For all statistical tests, p values of less than 0.05 were considered significant. Statistical package SPSS® version 22 was used for analyzes.

Results

Reliability of scale

The first source of reliability is a measure of internal consistency analyzed by Cronbach’s alpha. The result obtained is a = 0.97, showing a very high internal consistency.

The second is a measure of objectivity that helps us to determine how consistently judges evaluate the same cases using the scale. To do this, 5 judges were presented a total of 3 videos with fictitious cases. The evaluations were submitted to a Kendall coefficient of concordance W, obtaining a result of W = 0.96, a high and very significant concordance among judges (p <0.01).

The third is a test-retest measure among judges. Six judges evaluated 10 videos of real cases two months apart. The average Pearson correlation between the first and second evaluation was between 0.97 and 0.99 (table 1), indicating a high consistency among judges’ assessments.

Validity of the scale

The first evidence, related to validity of the scale content, is assured by the process of development of its items by experts, which was previously described in the section of Patients and method. In addition, content analysis was performed by a group of experts outside the team in charge of development.

The second source of validity, considered as evidence of concurrent validity, is the concordance of the scale results among judges, with the evaluation carried out by the evaluator 1, who is expert judge, regarded as the measurement pattern. When comparing the means of the 5 judges with this expert, a non-significant difference was obtained (F = 0.023, p > 0.5).

A third source of evidence, which supports concurrent validity, is obtained by comparing the results of the scale with IB, which showed an average score of 48
<table>
<thead>
<tr>
<th>Item</th>
<th>Componente de funcionalidad</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Proximal</td>
</tr>
<tr>
<td>Abducción hombros hasta la altura de los hombros (D)</td>
<td>0,91</td>
</tr>
<tr>
<td>Abducción hombros sobre la altura de los hombros (D)</td>
<td>0,95</td>
</tr>
<tr>
<td>Flexión de hombros hasta la altura de los hombros (D)</td>
<td>0,94</td>
</tr>
<tr>
<td>Flexión de hombros sobre la altura de los hombros (D)</td>
<td>0,95</td>
</tr>
<tr>
<td>Abducción hombros hasta la altura de los hombros (I)</td>
<td>0,91</td>
</tr>
<tr>
<td>Abducción hombros sobre la altura de los hombros (I)</td>
<td>0,95</td>
</tr>
<tr>
<td>Flexión de hombros hasta la altura de los hombros (I)</td>
<td>0,94</td>
</tr>
<tr>
<td>Flexión de hombros sobre la altura de los hombros (I)</td>
<td>0,95</td>
</tr>
<tr>
<td>Manos a la boca</td>
<td>0,28</td>
</tr>
<tr>
<td>Trasladar peso desde los muslos a la mesa o a la altura de los hombros con las 2 manos</td>
<td>0,35</td>
</tr>
<tr>
<td>Levantar y trasladar latas</td>
<td>0,25</td>
</tr>
<tr>
<td>Rasgar papel</td>
<td>0,45</td>
</tr>
<tr>
<td>Desplazar peso de un círculo a otro</td>
<td>0,41</td>
</tr>
<tr>
<td>Trazar trayecto en hoja</td>
<td>0,19</td>
</tr>
<tr>
<td>Encender la luz presionando el interruptor</td>
<td>–0,20</td>
</tr>
<tr>
<td>Agarrar 5 monedas</td>
<td>–0,01</td>
</tr>
<tr>
<td>Levanta con agarre de 3 puntos de apoyo (pinza trípode)</td>
<td>0,05</td>
</tr>
<tr>
<td>Levanta con agarre de 2 puntos de apoyo (pinza término-terminal)</td>
<td>0,09</td>
</tr>
<tr>
<td>Ponerse una camiseta</td>
<td>0,74</td>
</tr>
<tr>
<td>Llevar lata llena de bebida a la boca</td>
<td>0,50</td>
</tr>
<tr>
<td>Llevar cuchara a la boca</td>
<td>0,35</td>
</tr>
<tr>
<td>Peinarse</td>
<td>0,51</td>
</tr>
<tr>
<td>Lavarse los dientes</td>
<td>0,48</td>
</tr>
<tr>
<td>Abrir una botella</td>
<td>0,17</td>
</tr>
<tr>
<td>Abrir la tapa de un recipiente</td>
<td>0,06</td>
</tr>
</tbody>
</table>

Matriz rotada análisis factorial.
among the 10 patients (range 20 to 90). Although IB contains aspects not evaluated by our scale, the correlation was very high and significant ($r = 0.93$).

Finally, we have evidence of scale development validity through factor analysis (table 2). A Factorial analysis of the scale indicated 4 factors that explain 85.44% of the total variance. The first factor is proximal functionality, the second middle functionality, the third distal functionality and the fourth is of mixed functionality. The first one includes items of shoulder functionality and the process of putting on a T-shirt. The second one involves the processes of bringing hands to the mouth, transfer weight from the thighs to the table or to shoulder height with the two hands, lift and transfer cans, move weight from one circle to another, draw path on a paper, bring a full can to mouth, bring a spoon to the mouth, combing and brushing their teeth. The third one is the process of tearing a piece of paper and lifting an object with 2 point grip and opening the lid of a container. The fourth one consists of grabbing 5 coins and opening a bottle.

It is interesting to note that the most complex items on the scale, those that have significant factor weights in more than one factor, are precisely those that evaluate functionality of daily activities, such as putting on a T-shirt, bringing a full can to the mouth, bringing a spoon to the mouth, combing and brushing the teeth.

**Discussion and Conclusions**

This article describes an UL function scale evaluation in children with lack of strength secondary to 2 of the most frequent NMDs in pediatric age, DMD and SMA. The application of the scale requires a process of self-training and the use of inexpensive and easily acquired materials to create the stimuli used during the application. This scale proves to be highly reliable and shows a high concordance among and within the evaluators.

The need to have an UL function evaluation instrument in patients with progressive loss of muscle strength, especially in stages near or after the inability of walking, is due to the fact that these children present significant axial and lower extremity lack of muscle strength. This results in that the activities that these patients can perform the best are those that involve the use of their upper limbs, usually not considered in most of the existing scales of functionality.

In the selection, aspects that were not the exclusive expression of muscular strength, but reflected the ability to perform functional actions were included. We selected those tests already reported in the literature and which seemed to us more representative of strength-associated functionality. The scale developed showed excellent reliability, with a very high internal consistency (Cronbach’s of 0.97).

The fact that the loss of muscular strength and second motor neuron disorders is characterized by a proximal to distal progression should be considered when grouping the items of an evaluation scale for this type of diseases if differentiate degrees of strength expressed in functionality is intended. Mayhew et al. suggested grouping the items in 3 levels: high, middle and distal. However, it is important to consider that the functionality of specific actions is affected by compensatory strategies that each patient develops through the evolution of the disease, and it is not only altered by the lack of strength in certain body parts such as the shoulders, elbows and/or wrists. Therefore, we thought that it was important to consider functionality actions represented by basic ADLs involving different segments of the upper limbs. Our scale was built in 4 dimensions: proximal, middle, distal and mixed functionality.

The factorial analysis of the scale was valid not only when it showed that 4 factors explain 85% of total variance, but also when described the presence of items with significant factorial weights in more than one factor. All these complex items correspond to the mixed functionality dimension. On the other hand, the high scale correlation ($r = 0.97$) with another instrument widely used in the evaluation of ADLs in DMD, such as IB, provides further evidence of its validity. In addition, IB has shown an important floor effect when applied to very weak patients with DMD and poor motility, suggesting the need to use instruments capable of adequately evaluating functionality in patients with a high degree of functional compromise, something that our scale shows to be capable of doing.

Literature suggests that the use of functional assessment instruments, based on clinical observation, requires specific training of the evaluator to achieve adequate reliability and consistency. This aspect limits their use as training is not always readily available to the evaluators. Our scale, applied by self-trained evaluators, showed a high inter-evaluator agreement with a Kendall W coefficient of 0.96 ($p < 0.001$) and an average intra-judges correlation greater than 0.97, proving the effectiveness of the self-learning strategy. The high levels of reliability obtained in the application of the scale make possible to avoid the need for training for a correct application. Another advantage that presents the scale is the easy implementation of the battery used, formed by elements constructed with accessible and low cost materials. Finally, the scale is in Spanish, overcoming another limitation described in the literature regarding the lack of instruments developed in our language and the need of translation of instruments that have been validated in other languages.
The limited number of participants may be considered a weakness of the study, especially in the case of SMA. However, this does not detract from the findings, since there was no difference in the results obtained between those patients who maintained the ability to walk with respect to those who had lost it. There was also no difference between children with DMD compared to children with SMA. Therefore, this scale is able to provide objective information on UL functionality in these patients, even at different stages of the evolution of the disease. Future studies involving a greater number of patients, with DMD and SMA at different stages of the disease, and other types of myopathies with significant strength compromise, especially UL, are crucial.

In summary, this scale is a reliable and valid instrument to evaluate UL functionality in children with DMD and SMA between 9 and 19 years of age. In addition, it is easy to implement due to the possibility of self-training and the use of simple and inexpensive materials.

**Ethical Responsibilities**

**Protection of people and animals:** The authors reported that no experiments on either people or animals have been performed.

**Confidentiality of personal data:** The authors reported that they have followed the protocols of their center regarding the publication of personal data.

**Privacy rights and informed consent:** The authors have obtained the informed consent from patients and/ or subjects referred to in the article. These documents are in the possession of the corresponding author.

**Funding**

Project funded by Vice-Rectory for Research, UC Interdisciplinary Research Center, Pontifical Catholic University of Chile.

**Conflict of interests**

The authors declare no conflict of interest.

**Acknowledgements**

We thank the OT Rodolfo Aviles, Vanesa Bron, Miguel Jorquera, Loreto Quezada, Maria Paz Salinas y Bernardita Severin; all part of the staff at Neurorehabilitation and Neuromuscular Disease Laboratory, Catholic University, for your help to determine the validity and reliability of the scale.

**References**

17. Merlini L, Mazzone ES, Solari A, Morandi L. Reliability of handheld dynamometry...