RESEARCH ARTICLE

Development of a Functional Assessment Scale for Ambulatory Boys with Duchenne Muscular Dystrophy

Elaine Scott[1]*, Michelle Eagle[2], Anna Mayhew[2], Jenny Freeman[3], Marion Main[4], Jennie Sheehan[5], Adnan Manzur[6], Francesco Muntoni[6] & The North Star Clinical Network for Paediatric Neuromuscular Disease[†]

[1]Muscular Dystrophy Campaign, London, UK

[2]Institute of Human Genetics, Newcastle, UK

[3]University of Sheffield, Sheffield, UK

[4]Great Ormond Street Hospital, London, UK

[5]Evelina Children's Hospital, London, UK

[6]Dubowitz Neuromuscular Centre, Institute of Child Health, UCL, London, UK

Abstract

Background and Purpose. The aims of this study were to develop a clinical assessment scale to measure functional ability in ambulant boys with Duchenne muscular dystrophy and to determine the reliability of the scale in multiple centres in the UK. **Methods.** Focus groups and workshops were held with experienced paediatric neuromuscular physiotherapists to determine scale content. A manual was prepared with accompanying videos, and training sessions were conducted. A total of 17 physiotherapists from participating centres used the videos to determine inter-rater reliability. Five determined the intra-rater reliability. **Results.** Strength of agreement for these groups based on total subject scores was very good (0.95 and ≥0.93 for consistency and absolute agreement, respectively). Test–retest ability was high, with perfect agreement between occasions for all but two items of the scale. **Conclusions.** Our study indicates that the North Star Ambulatory Assessment is practical and reliable. It takes only 10 minutes to perform and incorporates both universally used timed tests as well as levels of activities, which allow assessment of high-functioning boys with Duchenne muscular dystrophy.

Received 27 January 2011; Revised 21 May 2011; Accepted 26 June 2011

Keywords

ambulant; assessment; Duchenne muscular dystrophy

***Correspondence**

Elaine Scott, MPhil, MCSP, c/o Muscular Dystrophy Campaign, 61 Southwark Street, London SE1 0HL, UK.

Email: elaines@muscular-dystrophy.org

[†]The North Star Clinical Network for Paediatric Neuromuscular Disease: Collaborators (53): Manzur A.Y., Muntoni F., Robb S., Main M., Kemp J. (Great Ormond Street Hospital, London), Scott E. (Muscular Dystrophy Campaign, London), Bushby K., Straub V., A. Sarkozy, E. Strehle, R. Venkateswaran, Eagle E., Mayhew A. (Institute of Human Genetics, Newcastle), Roper H., McMurchie H., Grace A. (Heartlands Hospital, Birmingham), Spinty S., Peachey G., Shillington S. (Alder Hey Children's Hospital, Liverpool), Quinlivan R., Groves L. (Robert Jones and Agnes Hunt Royal Orthopaedic Hospital, Oswestry), Wraige E., Jungbluth H., Sheehan J., Spahr R. (Evalina Children's Hospital, London), Hughes I., Bateman E., Cammiss C. (Royal Manchester Children's Hospital), Childs A.M., Pallant L., Psyden K. (Leeds General Infirmary), Baxter P. (Sheffield Children's Hospital), Naismith K., Keddie A. (King's Cross Hospital, Dundee), Horrocks I., McWilliam R., Di Marco M. (Yorkhill Children's Hospital, Glasgow), Hartley L., Sheen B., Fenton-May J. (University Hospital Wales, Cardiff), Jardine P., Majumdar A., Jenkins L. (Frenchay Hospital, Bristol), Chow G., Miah A. (Queen's Medical Centre University Hospital, Nottingham), de Goede C. (Preston Royal Hospital), Thomas N., Geary M.,

Palmer J (Southampton General Hospital), White C., Greenfield K. (Morriston Hospital, Swansea), MacAuley S. (Royal Belfast Hospital for Sick Children), Baxter A., Yirrell Y., Longman C. (Royal Hospital for Sick Children, Western General Hospital, Edinburgh).

Published online 23 September 2011 in Wiley Online Library (wileyonlinelibrary.com) DOI: 10.1002/pri.520

Introduction

The development of validated outcome measures in Duchenne muscular dystrophy (DMD) has become increasingly important because of the need to monitor disease progression and the impact of treatments, such as intermittent or daily steroids (Manzur et al., 2008a), and the requirement for reliable functional outcome measures for use in clinical trials (Manzur et al., 2008b; Mercuri et al., 2008). Since the original work of Brooke et al. (1981) and Scott et al. (1982) in the early 1980s on outcome measures for DMD, numerous papers relating to the use of a large variety of outcome measures for this condition have been published (e.g. McDonald, 2002; Kohler et al., 2005; Scott and Mawson, 2006; Mayhew et al., 2007; Davis et al., 2010; McDonald et al., 2010), and much work has been carried out, or is currently underway, to evaluate the usefulness of these measures in clinical and research settings via the TREAT-NMD clinical and research EU-funded network of excellence (www.treat-nmd.eu) and other related programmes. The focus of this study, however, is on the development and validation of a functional scale for needs of the North Star Project in the UK.

The North Star Clinical Network for Paediatric Neuromuscular Disease Management is a collaboration of 17 specialist neuromuscular centres in the UK whose overall aim is to optimize the management of children with DMD (Scott et al., 2007). The key objective of the network was to standardize clinical assessment protocols and pharmacological interventions such as corticosteroid use for ambulant boys with DMD. Clinicians and therapists from the clinical network identified functional measures as the most clinically relevant outcomes in monitoring disease progression (*North Star network internal report*), and patients and their relatives relate more to measures of function as opposed merely to muscle strength. There is a need for such functional measures to be valid for the patient group and context in which they are to be used, to provide reliable data, which in the context of the North Star Project means data from multiple centres, to be responsive to change and to be feasible for the patient group and setting.

Many scales have been developed previously to assess the functional abilities of boys with DMD and other neuromuscular conditions; however, each of these scales had limitations as evaluation instruments for steroid-treated ambulant boys with this condition. These included a lack of sensitivity to change and a lack of data on reliability, practicality and ease of use across multiple clinical centres. The scales developed specifically for DMD by Vignos et al. (1963) and Brooke et al. (1981) provide simple ordinal-level data, which do not offer the degree of sensitivity required to assess the effect of novel treatments. The Egen Klassifikation (EK) Scale (Steffensen et al., 2001, 2002) and the Motor Function Measure (Bérard et al., 2005, 2006) are both examples of substantial, comprehensively developed measures for the assessment of neuromuscular disorders. However, the EK Scale addresses non-ambulant function only, and the Motor Function Measure is lengthy and neither disease nor stage specific. The most commonly used functional scale in the UK, the Hammersmith Motor Ability Scale (HMAS) (Scott et al., 1982), is satisfactorily used within a clinical setting; but reliability and validity have never been established. Furthermore, the scale was developed before corticosteroids were widely used and suffers from ceiling effects in children benefitting from the positive effect of this medication (*personal observation of the authors*). The focus of this study is therefore on the initial development of a functional assessment scale for ambulant children with DMD, the North Star Ambulatory Assessment (NSAA).

There were two aims of this study:

- to describe the initial development of the scale as a clinical assessment tool
- to present the initial training and reliability data from the North Star group of physiotherapists.

Methods

Development of the North Star Ambulatory Assessment

Year 1 — construction and development

The intent was to develop a clinical scale to evaluate change in the physical abilities of ambulatory boys with DMD. This included boys across a range of ability levels

from those only just able to walk to those with higher functional abilities such as running and jumping. As there are 17 centres involved in the North Star clinical network, it was important that minimal equipment should be required for standardization of use across sites. Given the behavioural difficulties that are common in young boys with DMD and the limited time available in clinic settings, it was also important that the time to complete the scale should be kept to a minimum.

Although the HMAS has the limitations stated previously, it has an acknowledged clinical utility and a long history of use in the UK. It was therefore used as a basic framework from which the NSAA was developed. A focus group of specialist neuromuscular physiotherapists (M. M., M. E. and J. S., with E. S.) was convened to determine the structure and content of the scale. Domain of content (Portney and Watkins, 2000) for the NSAA was defined as the gross motor ability in ambulant DMD children. The underpinning theoretical construct is that these boys lose functional ambulation in a recognizable pattern due to the primary underlying pathology of progressive muscle deterioration and related complications such as contractures. Activities included were those necessary to remain functionally ambulant including the ability to rise from the floor and getting from sitting to standing. Head raise and standing on heels were included as these are difficult even in the early stages of the disease while children are still ambulant and were seen as clinically relevant. Other activities such as hopping, jumping and running are unusual in non-steroid-treated children yet are frequently seen in children treated with steroids, whether on intermittent or daily regimes. An example of items included in the NSAA is given in Table 1.

The scores for each item are described in terms of clinically significant changes in the functional abilities seen in this patient group, reflecting the pattern of disease progression. Sensitivity to change has thus been addressed theoretically with the description of the item categories in terms of 'clinically significant change' (Bain and Dollaghan, 1991), that is, change that denotes a true change in a patient's abilities not merely due to natural variability of performance, or maturation. Following a series of four focus group meetings, a manual to enable standardized use of the scale was developed and introduced at a series of eight workshops for experienced neuromuscular physiotherapists in the North Star network. After a six-month period of clinical use and assessment, the scale was formally reviewed, and amendments were made to the description of the activities to make it easier to grade each task; however, no new items were included, and none were removed. In this manner, face and content validity were addressed by the initial focus group then further validated by the wider expert group of the North Star

Table 1. North Star Ambulatory Assessment, example of test items

Test item 1: Stand	
Starting position	Feet should be no further than 10cm apart and heels on the ground if possible. Arms by sides. NO shoes should be worn.
Instruction	Can you stand up tall for me for as long as you can and as still as you can for three seconds with your heels flat on the ground?
Scoring detail	When counting to 3 – Use "**And 1 - and 2 - and 3**" so that three seconds is achieved on the word of 3. Best done on the floor rather than on a mat. Whichever is chosen maintain consistency through repeated testing sessions. Score 2 - Minimum count of 3 seconds.
Score	
2	Stands upright, still, symmetrical, without compensation (heels flat legs in neutral) for minimum count of 3 seconds
1	Stands still but with compensation (e.g. on toes or with legs abducted or with bottom stuck out) for minimum count of 3 seconds
0	Cannot stand still or independently, needs support (even minimal)
Test item 14: Jump	
Starting position	Standing on the floor, feet fairly close together. No shoes should be worn.
Instruction	How high can you jump?
Scoring detail	Want height, not forward movement. Small amount of forward movement acceptable
Score	
2	Both feet at the same time, clear the ground simultaneously
1	One foot after the other (skip) or does not fully clear both feet at the same time.
0	Unable

clinical network. Figure 1 summarizes the process of development and review of the measure.

Year 2 — reliability testing and ongoing training

Initial training in the use of the scale included workshops, centre visits by the project coordinator, where joint patient assessments were undertaken, and a period where assessors piloted the scale in clinical practice. Two sets of reliability data are presented:

- six subjects evaluated by five of the North Star group of therapists
- three subjects evaluated by 17 therapists from all participating centres.

Table 2 provides an outline as to which raters evaluated which subjects for these data sets. Video was taken of boys with a range of differing abilities, and each evaluator was asked to independently score each child. Scoring was carried out in confidence during group sessions, with the coordinator present. Video of scale items could be viewed more than once on request. Score sheets were submitted to the coordinator.

Further, to this intra-rater reliability was examined with five experienced physiotherapists who independently scored the same individual from video on two occasions, with a one-month interval between evaluations.

The scale has subsequently been adopted as a standard clinical assessment tool for use in over 17 paediatric neuromuscular centres in the UK.

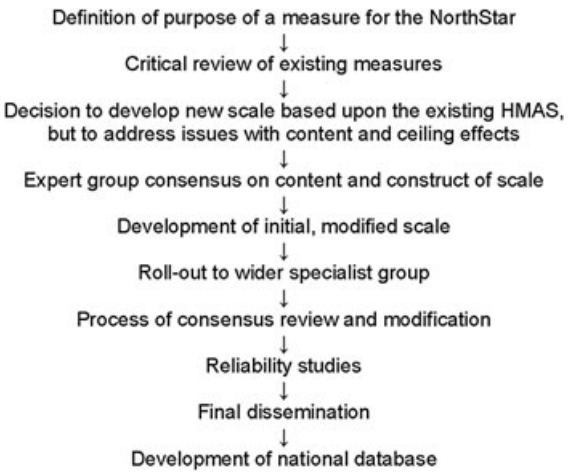

Figure 1 Summary of the North Star Ambulatory Assessment development process

Table 2. Outline of rater to subject evaluations

	Raters																
Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
A	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
B	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
C	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x
D	x	x	x	x	x												
E	x	x	x	x	x												
F	x	x	x	x	x												

Statistical analysis

Inter-rater reliability was assessed according to the methods outlined in the study by Streiner and Norman (2003). The intraclass correlation coefficient (ICC) for consistency among raters and the ICC for absolute agreement were computed. The ICC for consistency assesses whether raters were consistent in the order in which they placed individuals, that is, were boys rated in the same order for functional ability by all raters from worst to best (irrespective of the actual value). The ICC for absolute agreement assesses whether raters agreed with each other with respect to the actual values they assigned individuals. Given the mathematical equivalence between the ICC and the kappa statistic (Fleiss and Cohen, 1973), interpretation of the tabulated ICC values was based on the semantic categories adapted by Altman (1991) from Landis and Koch, as shown in Table 3. However, the kappa statistic provides a poor summary measure of agreement when prevalence is low, as it was for some of the features investigated here. As a result, the percentage classified into each category across all raters and all children ($n = 30$ for the smaller group of five therapists, and $n = 51$ for the group of 17 therapists) has also been tabulated for each variable. Intra-rater reliability was assessed using data from five assessors who each assess the same, single individual on two occasions. As agreement was perfect for all but two

Table 3. Interpretation of kappa statistic and intraclass correlation coefficient

Value of K	Strength of agreement
<0.20	Poor
0.21–0.40	Fair
0.41–0.60	Moderate
0.61–0.80	Good
0.81–1.00	Very good

of the items on the scale, it was not possible to calculate the test–retest coefficient, and thus, only the percentage agreement has been presented.

Results

Year 1 — construction and development of the scale

Following a substantial period of development and specialist review, a 17-item scale was agreed. Aspects of theoretical construct and content validity were addressed, as was the feasibility of use in multiple clinical centres for this patient group. A document standardizing the test method was compiled and circulated to all involved in the clinical network, and workshops and site visits were undertaken to ensure standardized application of the scale (full test details are available from www.muscular-dystrophy.org/how_we_help_you/for_professionals/clinical_databases).

Year 2

Five physiotherapists evaluated the videos of six boys by using the NSAA scale. Strength of agreement for this group (Table 4) based on total subject scores was very good (0.95 for both consistency and absolute agreement). Fifteen of the 17 individual items were rated good or very good for consistency, and 15 for absolute agreement. The 17 physiotherapists involved in data collection for the network evaluated the videos of three boys. All 17 evaluated all videos. As with the previous data for the five physiotherapists and six boys, the strength of agreement (Table 5) when based on total subject scores was very good (0.95 and 0.93 for consistency and absolute agreement, respectively). Nine of the individual items were rated good or very good for consistency, and nine for absolute agreement. Where items have been rated poor or fair on ICC analysis, for example, rise from floor, there was an overall good agreement among therapists by the percentage classified into category 1, that is, over 90% of the therapists classified the subjects as scoring 1 (signs of Gowers' manoeuvre). This, however, meant that there was a low prevalence across the range of scores — none of the subjects in this study were scored as unable to rise from the floor (Tables 4 and 5). This issue of prevalence across the range of scores (2,1,0) and its effect upon results will be returned to in the Discussion and Conclusion sections.

Although the numbers were small (only five physiotherapists), the agreement between occasion 1 and occasion 2 was perfect for all but two of the items, jump and run (Table 6).

Table 4. Inter-rater reliability: percentage classified into each category, together with intraclass correlation coefficients (ICCs), $n = 30$ (six subjects, five assessors)

	% classified as:			ICC:	
Test item	0	1	2	Consistency	Absolute agreement
Stand	3.3	46.7	50.0	0.91	0.90
Walk		53.3	46.7	0.89	0.87
Sit to stand		50.0	50.0	0.78	0.75
Single leg stand (right)		50.0	50.0	1.00	1.00
Single leg stand (left)		33.3	66.7	0.75	0.72
Climb step (right)		16.7	83.3	1.00	1.00
Climb step (left)		20.0	80.0	0.82	0.80
Descend step (right)		23.3	76.7	0.76	0.73
Descend step (left)		40.0	40.0	0.68	0.64
Lying to sitting[1]		55.2	44.8	0.41	0.38
Rise from floor		96.7	3.3	0.00	0.00
Lift head	66.7	23.3	76.7	0.76	0.73
Stand on heels	10.0	16.7	16.7	0.85	0.83
Jump	30.0	26.7	63.3	0.78	0.75
Hop (right)	36.7	30.0	40.0	0.76	0.73
Hop (left)	23.3	40.0	23.3	0.78	0.75
Run		26.7	50.0	0.74	0.71
Total score				0.95	0.95

[1]One observation missing.

Table 5. Inter-rater reliability: percentage classified into each category, together with intraclass correlation coefficients (ICCs), $n = 51$ (three subjects, 17 assessors)

	% classified as:			ICC:	
Test item	0	1	2	Consistency	Absolute agreement
Stand	3.9	60.8	35.3	0.87	0.83
Walk		66.7	33.3	0.88	0.84
Sit to stand		41.2	58.8	0.83	0.78
Single leg stand(right)		68.6	31.4	0.94	0.92
Single leg stand (left)		64.7	35.3	0.94	0.92
Climb step (right)		2.0	98.0	0.00	0.00
Climb step (left)		13.7	86.3	0.38	0.31
Descend step (right)		19.6	80.4	0.28	0.22
Descend step (left)		19.6	80.4	0.30	0.24
Lying to sitting		64.7	35.3	0.70	0.64
Rise from floor		90.2	9.8	0.10	0.08
Lift head		7.8	92.2	0.05	0.04
Stand on heels	66.7	23.5	9.8	0.81	0.77
Jump	21.6	25.5	52.9	0.60	0.53
Hop (right)	29.4	41.2	29.4	0.78	0.73
Hop (left)	39.2	47.1	13.7	0.74	0.68
Run	37.2	21.6	41.2	0.54	0.47
Total score				0.95	0.93

Table 6. Agreement between occasion 1 and occasion 2 ($n = 5$)

Test item	Agreement (%)
Stand	100
Walk	100
Sit to stand	100
Single leg stand (right)	100
Single leg stand (left)	100
Climb step (right)	100
Climb step (left)	100
Descend step (right)	100
Descend step (left)	100
Lying to sitting	100
Rise from floor	100
Lift head	100
Stand on heels	100
Jump	80
Hop (right)	100
Hop (left)	100
Run	60

Discussion

The NSAA has been developed by expert paediatric neuromuscular physiotherapists specifically for use in ambulant children with DMD, thus ensuring that the content of the scale is clinically meaningful and appropriate. Experience from clinical use has shown that the NSAA takes approximately 10 minutes to complete, including timed tests, and its ease of administration means that it can be used both in specialist clinics and community settings. Evaluation of the feasibility of its administration by the North Star therapists indicates that compliance is good even in children with learning or behavioural problems, a feature that characterizes one-third of all DMD boys (Emery and Muntoni, 2003).

Lead physiotherapists from all 17 centres involved in data collection for the project participated in the reliability exercise, where the main focus was on ensuring standardized scoring and rater agreement. Strength of agreement when based on total subject scores was found to be excellent. This varied substantially, however, when individual test items were considered, particularly for the larger therapist group. Although good to very good agreement on analysis by ICC was gained for many test items on the NSAA, very poor results were gained for six by the larger therapist group. This apparently poor strength of agreement is confounded when the percentage classified into each category is considered. The percentage classified results show that there was actually excellent agreement among therapists for most of these test items, but the majority of subjects were considered to fall into one category (see item 11 (rise from floor) in Tables 3 and 4). The kappa statistic, and therefore the ICC, provides a poor summary measure of agreement when prevalence of

different categories is low, as it was for some of the features investigated here, and so it should be interpreted with caution. Although the three subjects presented with varying abilities, all achieved an average total score of 18 or over. The evaluation of a bigger group of subjects with a wider variety of abilities would have addressed these issues; however, even this may not alleviate this issue where prevalence in any one category is low (Fleiss and Cohen, 1973).

A subsequent study in Italy by Mazzone et al. (2009) reported their findings regarding training needs and reliability studies for the NSAA for a substantially larger group of patients. Their initial results, following the first phase of training, were poor (ICC <0.05), but following a second phase, the ICC for all items was ≥ 0.75, with all but one item indicating a very good level of agreement. Two of the items (lift head and run), which proved problematic to gain agreement in the first phase of the Italian study, were also the same for the larger North Star group (Table 4). However, without the percentage classified data, it is not possible to know if these poor ICC results are an artefact of the low prevalence across categories or of a lack of clarity with the wording of the scale. Neither of these items were a cause for concern with the smaller North Star group or the Italian group following their second phase of training. The issue of training and in particular consensus building, as demonstrated in Mazzone's paper, is an important one for any clinician-rated assessment scale, such as the NSAA. The medium that is used to translate the patient's performance of an activity to a point on the scale is observation on the part of the assessor. The assessor's interpretation/understanding of the wording of the scale therefore becomes an important factor in avoiding measurement error and achieving reliable results from the scale. Translation into another language may also play a factor in the interpretation of the scale and needs careful consideration in international studies. Ongoing training for the North Star clinical network includes an annual reliability review to ensure consistency and quality of data collated for the national database. The Italian study reported excellent inter-rater reliability (ICC = 0.995) based on total scores comparable with the initial reliability studies from the North Star group. They also reported a very high level of intra-rater reliability (0.95) in their comprehensive work.

Although the initial phase of both the UK and the Italian studies indicated potential issues with reliability indices for a small number of the scale items, the decision was made, because of expert clinical opinion as to their clinical importance for this patient group, to continue to include these items. Further, rigorous evaluation of the psychometric properties of the scale in a large population of boys with DMD is currently underway using Rasch methodology. Following this, recommendations may be made for scale modifications, balancing high levels of validity and reliability with the need for clinical and statistical relevance.

The scale has now been in use for more than four years, and data from over 300 DMD boys are being collated, with the formal consent of families and assent of patients, on the North Star database in the UK. The database holds comprehensive national data on clinical performance and outcomes from patients who attend participating centres. Consistent and standardized longitudinal clinical data, of which the NSAA is part, are therefore being collated on a cohort of children with DMD offering a valuable tool for clinical audit and research purposes.

The North Star clinical scale is now also been used to document clinical response in experimental clinical trials (Cirak et al., 2009; Kinali et al., 2009). The new challenge, which is being considered, is how to expand the scale to include non-ambulant children and young adults, as natural history data in this group of individuals are scant.

Conclusion

The NSAA is a reliable, robust, practical test that can be realistically used across a range of settings with minimal, universally available equipment and completed in a short timeframe. The initial results of the reliability analyses for the North Star group of physiotherapy assessors show an excellent level of agreement, which is further demonstrated in the Italian study (Mazzone et al., 2009). The scale incorporates the important disease milestones such as rising from the floor and walking ability as well as incorporating new skills that are acquired by DMD boys treated with steroids. A study is currently underway to further evaluate the theoretical content and construct validity, reliability and sensitivity to change of the scale by using the Rasch methodology. The authors also plan to present the longitudinal data from the North Star database in due course.

Acknowledgements

The North Star Project is supported and funded by the Muscular Dystrophy Campaign. The MRC Neuromuscular Centre's support to the North Star Database and the Muscular Dystrophy Campaign Centre's grant to the Dubowitz Neuromuscular Centre are also gratefully acknowledged. Professor Muntoni is supported by the Great Ormond Street Hospital Children's Charity.

REFERENCES

Altman DG. Practical Statistics for Medical Research. London: Chapman & Hall, 1991.

Bain BA, Dollaghan CA. Clinical forum: treatment efficacy. The notion of clinically significant change. Language, Speech, and Hearing Services in Schools 1991; 22(10): 264–270.

Bérard C, Payan C, Fermanian J, Girardot F, Groupe d'Etude MFM. A motor function measurement scale for neuromuscular diseases—description and validation study (in French). Revue Neurologique (Paris) 2006; 162(4): 485–493.

Bérard C, Payan C, Hodgkinson I, Fermanian J, MFM Collaborative Study Group. A motor function measure for neuromuscular diseases. Construction and validation study. Neuromuscular Disorders 2005; 15(7): 463–470.

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): 186–197.

Cirak S, Muntoni F, Guglieri M, Bushby K, Leow S, Shrewsbury S. Current progress with the systemic administration trial of AVI-4658, a novel phosphorodiamidate morpholino oligomer (PMO) skipping exon 51 in Duchenne muscular dystrophy (DMD). World Muscle Society Meeting, Geneva, 9–12 September 2009. Neuromuscular Disorders 2009; 19(8–9): 511–674.

Davis SE, Hynan LS, Limbers CA, Andersen CM, Greene MC, Varni JW, Iannaccone ST. The PedsQL in pediatric patients with Duchenne muscular dystrophy: feasibility, reliability, and validity of the Pediatric Quality of Life Inventory Neuromuscular Module and Generic Core Scales. Journal of Clinical Neuromuscular Disease 2010; 11(3): 97–109.

Emery EH, Muntoni F. In: Emery AEH, Muntoni F (eds), Duchenne Muscular Dystrophy (3rd edition). Oxford: Oxford University Press, 2003.

Fleiss JL, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. Educational and Psychological Measurement 1973; 33: 613–619.

Kinali M, Arechavala-Gomeza V, Feng L, Cirak S, Hunt D, Adkin C, Guglieri M, Ashton E, Abbs S, Nihoyannopoulos P, Garralda ME, Rutherford M, McCulley C, Popplewell L, Graham IR, Dickson G, Wood MJ, Wells DJ, Wilton SD, Kole R, Straub V, Bushby K, Sewry C, Morgan JE, Muntoni F. Local restoration of dystrophin expression in Duchenne muscular dystrophy: a single blind, placebo-controlled dose escalation study using morpholino antisense oligomer AVI-4658. Lancet Neurology 2009; 8(10): 918–928.

Kohler M, Clarenbach CF, Böni L, Brack T, Russi EW, Bloch KE. Quality of life, physical disability, and respiratory impairment in Duchenne muscular dystrophy. American Journal of Respiratory and Critical Care Medicine 2005; 172(8): 1032–1036.

Manzur AY, Kuntzer T, Pike M, Swan AV. Glucocorticoid corticosteroids for Duchenne muscular dystrophy. Cochrane Database of Systematic Reviews 2008a, Issue 1. Art. No. CD003725. DOI: 10.1002/14651858.CD003725.pub3

Manzur AY, Kinali M, Muntoni F. Update on the management of Duchenne muscular dystrophy. Archives of Disease in Childhood 2008b; 93(11): 986–990. Invited review.

Mayhew JE, Florence JM, Mayhew TP, Henricson EK, Leshner RT, McCarter RJ, Escolar DM. Reliable surrogate outcome measures in multicenter clinical trials of Duchenne muscular dystrophy. Muscle & Nerve 2007; 35(1): 36–42.

Mazzone ES, Messina S, Vasco G, Main M, Eagle M, D'Amico A, Doglio L, Politano L, Cavallaro F, Frosini S, Bello L, Magri F, Corlatti A, Zucchini E, Brancalion B, Rossi F, Ferretti M, Motta MG, Cecio MR, Berardinelli A, Alfieri P, Mongini T, Pini A, Astrea G, Battini R, Comi G, Pegoraro E, Morandi L, Pane M, Angelini C, Bruno C, Villanova M, Vita G, Donati MA, Bertini E, Mercuri E. Reliability of the North Star Ambulatory Assessment in a multicentric setting. Neuromuscular Disorders 2009; 19(7): 458–461.

McDonald CM, Henricson EK, Han JJ, Abresch RT, Nicorici A, Elfring GL, Atkinson L, Reha A, Hirawat S, Miller LL. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle & Nerve 2010; 41(4): 500–510.

McDonald CM. Physical activity, health impairments, and disability in neuromuscular disease. American Journal of Physical Medicine & Rehabilitation 2002; 81(11 Suppl): S108–S120.

Mercuri E, Mayhew A, Muntoni F, Messina S, Straub V, Van Ommen GJ, Voit T, Bertini E, Bushby K. TREAT-NMD Neuromuscular Network. Towards harmonisation of outcome measures for DMD and SMA within TREAT-NMD; report of three expert workshops: TREAT-NMD/ENMC workshop on outcome measures, 12th–13th May

2007, Naarden, The Netherlands; TREAT-NMD workshop on outcome measures in experimental trials for DMD, 30th June–1st July, Naarden, The Netherlands; Conjoint Institute of Myology TREAT-NMD meeting on physical activity monitoring in neuromuscular disorders, 11th July 2007, Paris, France. Neuromuscular Disorders 2008; 18: 894–903.

Portney L, Watkins M. Validity of Measurements. In: Foundations of Clinical Research. Applications to Practice. New Jersey: Prentice Hall Health, 2000: 79–110.

Scott E, Mawson SJ. Measurement in Duchenne muscular dystrophy: considerations in the development of a neuromuscular assessment tool. Developmental Medicine and Child Neurology 2006; 48(6): 540–544.

Scott E, Mayhew A, Muntoni F, Manzur AY. NorthStar and SMARTnet: development of paediatric neuromuscular disease networks in the UK. Developmental Medicine and Child Neurology 2007; 49(Supplement No. 108): 29.

Scott OM, Hyde SA, Goddard C, Dubowitz, V. Quantitation of muscle function in children: a prospective study in Duchenne muscular dystrophy. Muscle & Nerve 1982; 5(4): 291–301.

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. Physiotherapy Research International 2001; 6(3): 119–134.

Steffensen BF, Lyager S, Werge B, Rahbek J, Mattson E. Physical capacity in non-ambulatory people with Duchenne muscular dystrophy or spinal muscular atrophy: a longitudinal study. Developmental Medicine and Child Neurology 2002; 44(9): 623–632.

Streiner DL, Norman GR. 2003. Health Measurement Scales: A Practical Guide to Their Development and Use (3rd edition). Oxford: Oxford Medical Publications.

Vignos PJ, Jr, Spencer GE, Jr, Archibald KC. Management of progressive muscular dystrophy in childhood. JAMA 1963; 184: 89–96.