

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/309099222>

Revised upper limb module for spinal muscular atrophy: Development of a new module: RULM in SMA

Article in *Muscle & Nerve* · October 2016

DOI: 10.1002/mus.25430

CITATIONS

190

READS

11,983

26 authors, including:



[Anna Mayhew](#)

Newcastle University

205 PUBLICATIONS 3,431 CITATIONS

[SEE PROFILE](#)



[Jacqueline Montes](#)

Columbia University

168 PUBLICATIONS 10,330 CITATIONS

[SEE PROFILE](#)



[Danielle Ramsey](#)

University College London

29 PUBLICATIONS 689 CITATIONS

[SEE PROFILE](#)



[Lavinia Fanelli](#)

Policlinico Universitario Agostino Gemelli

51 PUBLICATIONS 1,366 CITATIONS

[SEE PROFILE](#)

REVISED UPPER LIMB MODULE FOR SPINAL MUSCULAR ATROPHY: DEVELOPMENT OF A NEW MODULE

ELENA S. MAZZONE, PT,^{1*} ANNA MAYHEW, PT,^{2*} JACQUELINE MONTES, PT,^{3*} DANIELLE RAMSEY, PT,⁴ LAVINIA FANELLI, PT,¹ SALLY DUNAWAY YOUNG, PT,³ RACHEL SALAZAR, PT,³ ROBERTO DE SANCTIS, PT,¹ AMY PASTERNAK, PT,⁵ ALLAN GLANZMAN, PT,⁶ GIORGIA CORATTI, PT,¹ MATTHEW CIVITELLO, PT,⁷ NICOLA FORCINA, PT,¹ RICHARD GEE, PT,⁸ TINA DUONG, PT,⁸ MARIKA PANE, MD,¹ MARIACRISTINA SCOTO, MD,⁴ MARIA CARMELA PERA, MD,¹ SONIA MESSINA, MD,⁹ GIHAN TENNEKON, MD,⁶ JOHN W. DAY, MD,⁸ BASIL T. DARRAS, MD,⁵ DARRYL C. DE VIVO, MD,³ RICHARD FINKEL, MD,^{7**} FRANCESCO MUNTONI, MD,^{4**} and EUGENIO MERCURI, MD^{1**}

¹Department of Paediatric Neurology, Catholic University, Largo Gemelli 8, 00168 Rome, Italy

²Institute of Genetic Medicine, Newcastle University, Newcastle, United Kingdom

³Department of Neurology, Columbia University Medical Center, New York, USA

⁴Dubowitz Neuromuscular Centre, UCL Institute of Child Health & Great Ormond Street Hospital, London, United Kingdom

⁵Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, USA

⁶Division of Neurology, Children's Hospital of Philadelphia and Departments of Neurology and Pediatrics, Perelman School of Medicine at the University of Pennsylvania, Philadelphia, USA

⁷Division of Neurology, Nemours Children's Hospital, Orlando, Florida, USA

⁸Department of Neurology, Stanford University, Stanford, California, USA

⁹Department of Clinical and Experimental Medicine and Nemo Sud Clinical Centre, University of Messina, Italy

Accepted 3 October 2016

ABSTRACT: *Introduction:* There is a growing need for a robust clinical measure to assess upper limb motor function in spinal muscular atrophy (SMA), as the available scales lack sensitivity at the extremes of the clinical spectrum. We report the development of the Revised Upper Limb Module (RULM), an assessment specifically designed for upper limb function in SMA patients. *Methods:* An international panel with specific neuromuscular expertise performed a thorough review of scales currently available to assess upper limb function in SMA. This review facilitated a revision of the existing upper limb function scales to make a more robust clinical scale. *Results:* Multiple revisions of the scale included statistical analysis and captured clinically relevant changes to fulfill requirements by regulators and advocacy groups. *Conclusions:* The resulting RULM scale shows good reliability and validity, making it a suitable tool to assess upper extremity function in the SMA population for multi-center clinical research.

Muscle Nerve 000:000–000, 2016

Additional supporting information may be found in the online version of this article.

Abbreviations: CI, confidence interval; DMD, Duchenne muscular dystrophy; ICC, intra-class correlation coefficients; NMD, neuromuscular disorder; PSI, Person Separation Index; PT, physical therapists; PUL, performance of upper limb; RULM, Revised Upper Limb Module; SMA, spinal muscular atrophy; UK, United Kingdom; ULM, Upper Limb Module; US, United States

Key words: functional assessment; Hammersmith functional motor scale; outcome measures; Rasch analysis; spinal muscular atrophy; upper limb function

*Authors Mazzone, Mayhew, and Montes, contributed equally to this work.

**Drs. Finke, Muntoni, and Mercuri are all senior authors.

Funding: Funding from the SMA Foundation to the PNCR network, from the SMA TRUST to the Dubowitz Neuromuscular Centre (D.R., M.S., and F.M.) for the SMA REACH UK Network, and from the Italian Telethon and Famiglie SMA to the Italian SMA network is gratefully acknowledged. F.M. also wishes to acknowledge the funding of the Muscular Dystrophy UK Centre grant; of the Medical Research Council Neuromuscular Centre; and the support of the National Institute for Health Research Biomedical Research Centre at Great Ormond Street Hospital for Children NHS Foundation Trust and University College London.

Correspondence to: E. Mercuri; e-mail: mercuri@rm.unicatt.it

© 2016 Wiley Periodicals, Inc.
Published online 00 Month 2016 in Wiley Online Library (wileyonlinelibrary.com). DOI 10.1002/mus.25430

Over the last few years there has been increasing attention to identification of suitable outcome measures for clinical trials in spinal muscular atrophy (SMA).^{1–4} Several measures of gross motor performance have been identified^{2,4} but none of the existing scales effectively capture the progressive muscle weakness representing the entire spectrum of the disease. This was particularly evident when assessing patients at the weak end of the spectrum with low gross motor function scores. Furthermore, there was a need to identify measures that were useful in young children, as existing scales often were validated above age 6 years and included items for which performance was age-dependent.

The original Upper Limb Module (ULM)⁵ was developed as an international effort by clinicians, physical therapists (PTs), researchers, and patient advocacy groups in an attempt to address these shortcomings. The original module was designed specifically to assess upper limb function in nonambulatory SMA patients, mainly targeting young children. The module used items from existing scales and newer items suitable for a wide age range, from 30 months to adults. The ULM was validated in a multicenter setting⁵ and has been used in SMA natural history studies^{6,7} and in ongoing clinical trials. It has proven very successful in assessing the originally targeted population, i.e., nonambulatory young children and weaker patients who have a floor effect or very low score on the Hammersmith Functional Motor Scale,⁵ significant lower limb contractures, and limited upper limb activities.

Nevertheless, critical analysis of the results indicated that a ceiling effect was present when assessing cohorts that include both weak and strong

nonambulatory and ambulatory patients, both in natural history studies and in clinical trials.^{6,7}

To address the ceiling effect and make the module more useful in a wider SMA population, we undertook a revision process and now report the steps leading to the development of the revised version of the ULM, called RULM.

MATERIALS AND METHODS

This study was performed as part of the collaboration between 3 networks based in the United States (US, the Pediatric Neuromuscular Clinical Research Network for SMA), the United Kingdom (UK, SMA Reach) and Italy (Italian SMA Network). The study was approved by the Ethics Committee or institutional research board at each center.

The development of the RULM followed several steps, including a review of upper limb clinical scales currently used in neuromuscular disorders (NMDs) as well as using clinical expertise and modern psychometric methods. Rasch analysis, already used in SMA⁸ can identify the strengths and weaknesses of rating scales⁹ while at the same time ensuring that scales fulfill current traditional psychometric criteria.¹⁰ This will provide data that are more clinically meaningful and provide detailed recommendations for scale modification. Further details on this methodology are described elsewhere.¹¹

Critical Review and Rasch Analysis of the Original ULM and Identification of Possible Additional Items. An expert panel consisting of clinicians and physical therapists (PTs) with specific neuromuscular knowledge and expertise in outcome measures and in SMA met to develop a conceptual framework reflecting the progression of weakness and natural history of functional decline in the upper limbs of both nonambulatory and ambulatory individuals with SMA.

The original version of the ULM⁵ was reviewed critically. Clinical expertise, natural history studies,^{6,7} and its use in clinical trials suggested that some items did not fully reflect the potential for change in more able patients and that the scoring system might benefit from some degree of adaptation. For these items, instead of 0–1–2 scores, we tried to add more categories that would indicate disease progression and better define scoring options.

To address the ceiling effect, new items were identified through a systematic and critical review of the existing upper limb scales in other NMDs,¹² and in particular the performance of upper limb (PUL) scale for Duchenne muscular dystrophy (DMD),¹³ a module that assesses a wider array of upper limb performance in DMD. A small suitability study was performed as part of an international

multicenter effort (Rome, London, Newcastle) using the PUL in 53 consecutive ambulatory and nonambulatory SMA patients (34 type 2; 19 type 3; age range, 2.3–23.4 years; mean, 9.5 years) to identify possible additional items that could also be suitable to assess the patterns of weakness observed in SMA. A new pilot version, ULM version 2 (ULM2), was developed taking into account the critical review and Rasch analysis of the original module (ULM), with the selection of the new items.

Revision of the ULM2 and Development of the RULM. The ULM2 was piloted in a multicenter setting (Rome, London, Newcastle) on 42 nonambulatory and ambulatory SMA patients of age between 3 and 71 years of age (median, 10 years) and compared with the original ULM using Rasch analysis. The results were further reviewed by an expert panel, highlighting possible difficulties in using the adapted wording of item instructions and scoring system in the new items. Rasch analysis showed generally good item fit, and improved targeting of patients when using the RULM compared with the ULM, although a small ceiling effect was still noted and there were some issues of disordered thresholds. This is related to response option ordering and indicates that scoring options were potentially too complex. Additional, more difficult items were added, and existing items were adapted to further reduce the ceiling effect. We also decided to simplify some of the scoring, as the widening of the scoring options introduced in the ULM2 did not always appear to work better than the original 0–2 scoring system. Minor modifications of wording and test equipment were also performed to facilitate standardization and clarity for better use in a multicenter setting.

The resulting RULM reflects these changes (Table 1). The RULM now has a total of 20 items with an entry item that serves as functional class identification and does not contribute to the total score. The remaining 19 scorable items reflect different functional domains and are graded on a 3-point system with a score of 0 (unable), 1 (able, with modification), and a maximum of 2 (able, no difficulty). There is only 1 item (item I) that is scored as a can/cannot score, with 1 as the highest score. The maximum total score is 37 (Supplementary Fig. S1, available online). A scoring sheet and manual with detailed graphics that provide better, standardized administration procedures related to general testing guidelines, patient positioning, equipment, item instructions, start and stop positions, and detailed scoring options were a part of the RULM revision. Assessments were performed

Table 1. Details of the items included in the final version of the Revised Upper Limb Module

Entry item
Bring hands from lap to table
Complete the path bringing the car to the finish line without stopping or taking pencil off of paper?
Pick up coins/tokens
Place coin/token into cup : On table: horizontal At shoulder height: vertical
Reach to the side and touch the coin/token: Bring hand at shoulder height and above
Push button light with one hand
Tearing paper
Open Ziploc container
Raise 200-g cup to mouth
Lift 200-g weight and bring it from 1 circle to the other (midline to outer circle on tested side) without sliding
Lift 500-g weight and bring it from 1 circle to the other (midline to outer circle on tested side) without sliding
Lift 200-g weight and bring it from one circle to the other (inner to outer circle on opposite side) without sliding across midline
Bring 500-g sand weight from lap to table or eye level
Bring both arms above head - <i>Shoulder abduction</i>
Bring 500-g weight above shoulder height - <i>Shoulder abduction</i>
Bring 1-kg weight above shoulder height - <i>Shoulder abduction</i>
Bring hand above shoulder height - <i>Shoulder flexion</i>
Bring 500-g weight above shoulder height - <i>Shoulder flexion</i>
Bring 1-kg weight above shoulder height - <i>Shoulder flexion</i>

unilaterally on the preferred arm chosen by the individual.

Refining the RULM. The final RULM version was used for wider circulation with training sessions among the 3 networks participating in the development of the revised module. An inter-network training session was performed for all participating therapists during a face-to-face meeting followed by video review to establish inter-rater reliability. Three (3) video assessments of the RULM were evaluated by 17 PTs across the 3 International Clinical Networks. Intra-class correlation coefficients (ICC) and 95% confidence intervals (CI) using a 2-way random effects, single measures, analysis of variance model to determine inter-rater reliability. The ICC (2,1) was 0.928 (95% CI, 0.761–0.998) for all 17 PT evaluators, (0.933 in the UK [$n=3$], 0.920 [$n=11$] in the US, and 0.947 ($n=3$) in Italy). Following the training sessions and implementation of assessments with the 3 networks, additional Rasch analysis was performed on a prospective international dataset including 134 ambulatory and nonambulatory patients ages 2 to 52 years (median, 9 years).

RESULTS

Overall, the RULM performed well against the battery of psychometric tests with minimal issues. The key findings from these analyses were as follows:

Fit. The item locations ranged from -5.19 to + 6.11, indicating that the RULM defines a good continuum with little overlap (very few items measure the same level of ability). The overall item–trait interaction χ^2 value was 183.99 (44 degrees of freedom). None of the 20 items had a fit residual outside the recommended range (-2.50 to + 2.50). Three items had a significant χ^2 probability ($P=0.001$; Table 2).

Targeting. The range of the upper histogram horizontally that does not extend outside the range of the lower histogram suggests no floor or ceiling effect of the current items unless individuals on the extreme ends of the spectrum are included where a small ceiling is noted (Fig. 1). The spread of items suggests a small amount of redundancy, with minimal gaps in measurement accuracy.

Response Categories. The individual item scoring threshold map in Figure 2 demonstrates no disordered items (0/20) even for the entry item A.

Reliability. RULM reliability was supported by a high Person Separation Index (PSI) (0.954).

Dependency. Only 2 pairs of items (shoulder abduction and flexion, and shoulder flexion and abduction with 500 g) had residuals that were highly correlated (>0.40) when excluding the entry item (not part of the total score), suggesting that a response to 1 influenced the response to the other. When 2 of these items were removed (shoulder abduction items) reliability (PSI) was not impacted (0.949) suggesting reliability was not artificially inflated.

Stability. There was no uniform or nonuniform differential item functioning for gender, indicating that gender did not influence performance on this scale.

The results of this analysis (Figs. 1 and 2) confirmed the robustness of the scale. The Rasch results were discussed, with feedback provided, by the evaluating clinicians. The scale was safely and easily administered with no refusals noted even in younger children. Key to this is that test duration never exceeded 20 min (range, 5–20 min).

The RULM was well-tolerated, and no muscle or joint injury occurred. A manual and a scoring sheet (Supplementary Fig. S1) provided clear instructions and scoring options for clinicians to follow, and, given training, it has been demonstrated that inter-rater reliability can be achieved (ICC always >0.9) Equipment was well sourced and standardized across all centers internationally. There was a reduced ceiling effect compared with the ULM, and each item could be associated with meaningful activities of daily living by patients.

Table 2. Individual item fit for the Revised Upper Limb Module in item location order (i.e., order of difficulty, easiest to most difficult)

Item N.	Item name	Location	SE	FitResid	ChiSq	P-value
2	Hands lap to table	-5.193	0.374	-0.207	1.081	0.582457
5	Pick up tokens	-3.702	0.285	0.189	14.527	0.000701*
1	Entry item	-2.905	0.121	-0.028	5.554	0.062239
3	Path	-2.882	0.246	0.341	12.642	0.001799*
19	Lift 200 g	-2.264	0.211	-0.369	4.894	0.086546
17	Cup to mouth	-2.216	0.209	-0.333	2.37	0.305737
6	Token in cup	-2.058	0.204	-0.491	0.858	0.651123
21	Lift 200 g diagonally	-1.335	0.183	-0.879	9.761	0.007593*
9	Reach to side	-0.515	0.171	-0.637	2.829	0.243071
20	Lift 500 g	-0.443	0.166	-0.23	1.102	0.576348
22	Lift weight from lap	-0.107	0.195	-0.889	3.436	0.179469
13	Tear paper	1.035	0.185	0.922	6.725	0.034658
24	Shoulder abduction	1.075	0.178	-0.256	1.075	0.584255
11	Push on light	1.19	0.193	1.558	9.017	0.011016
27	Shoulder flexion	1.219	0.186	-0.662	0.793	0.6726
14	Open ziploc	1.6	0.278	0.757	7.043	0.029549
25	Shoulder abduction 500 g	2.979	0.246	-0.634	4.642	0.098157
28	Shoulder flexion 500 g	3.555	0.258	-0.569	5.042	0.080396
26	Shoulder abduction 1 kg	4.857	0.312	-0.381	3.263	0.195599
29	Shoulder flexion 1 kg	6.108	0.381	-0.22	0.782	0.6763

*Items that have a significant χ^2 probability ($P = 0.001$).

There was full consensus from the expert clinicians that the RULM is the final most robust scale for assessment of upper limb function in SMA.

An illustration of the relationship of the original ULM to the RULM is included (Fig. 3), which also illustrates how a score on 1 scale can be predicted on another.

DISCUSSION

The development of the RULM described here reflects how the field of outcome measures

continues to evolve. In the past, most outcome measures were developed by adapting clinical tools that originated in a clinical setting, but these measures were not always suitable for research purposes.⁸ Other measures, in contrast, were developed using appropriate statistical methods but did not always capture the whole spectrum of the disease course or clinically relevant changes. More recently, regulators and advocacy groups have highlighted the need to have tools with items that reflect activities of daily living and are

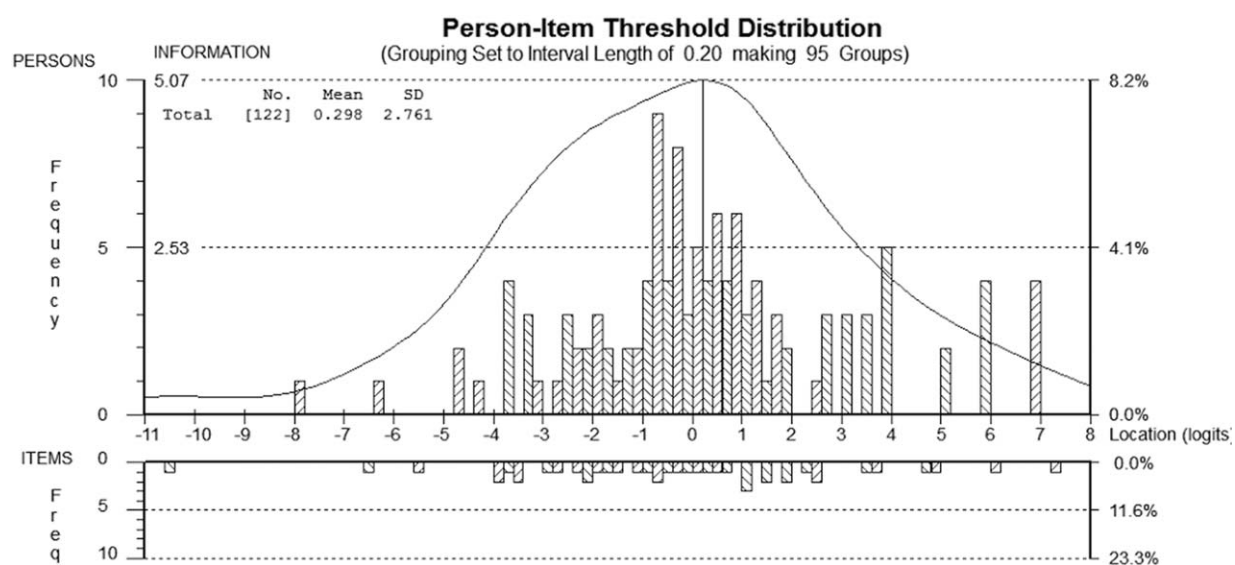


FIGURE 1. Person-item location distribution RULM (excludes extremes). Targeting of the patient sample (top) to the items (bottom). The figure shows good targeting between the distribution of person measurements (upper histogram) and the distribution of item locations (lower histogram). There are no ceiling/floor effects, as the range of the person measurements (upper histogram “blocks”) fits within the item locations (lower histogram “blocks”).

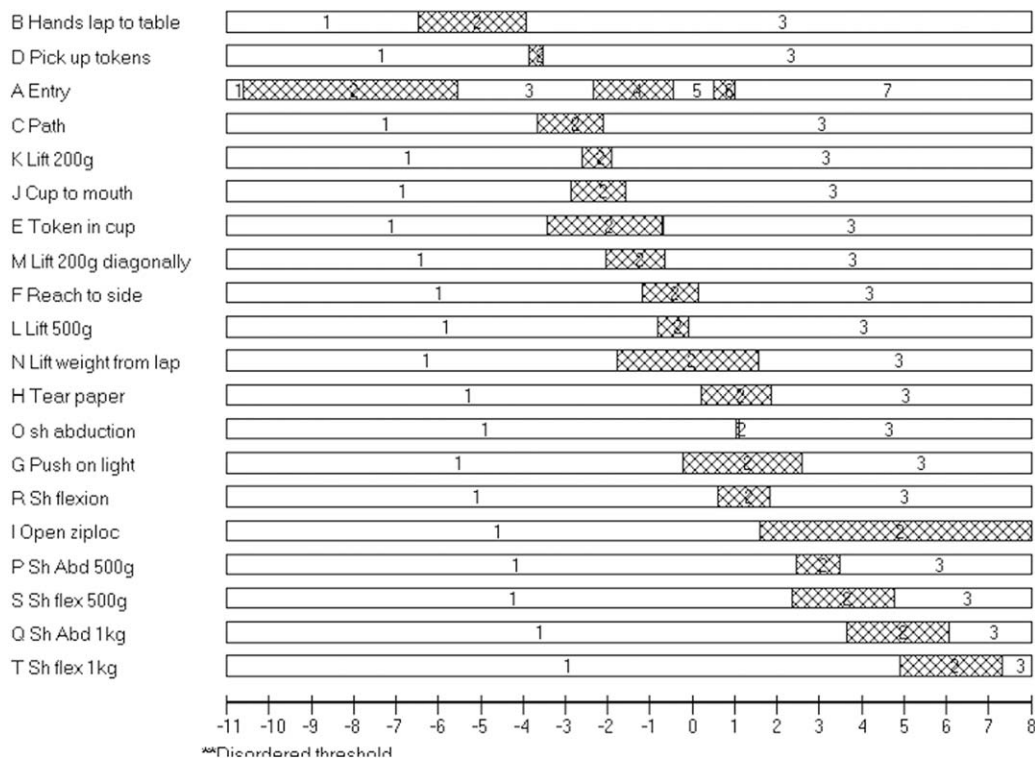


FIGURE 2. Threshold map for RULM items in ranked order of difficulty according to Rasch analysis. RULM has no disordered thresholds. 1 = Score of 0, unable; 2 = Score of 1, with modification; 3 = score 2, able, no difficulty. Item scoring 1–7 equivalent to entry item, 0–6.

“clinically meaningful” to patients and caregivers.^{4,14} The revision process addressed these bottlenecks by integrating all the different approaches.

We used a stepwise approach, starting from the original ULM, a module that was devised specifically for SMA and focused on activities observed in younger and weaker SMA patients. The RULM fits the conceptual framework for SMA with an

increased number of items and captures a wider spectrum of functional abilities across ambulatory and nonambulatory individuals with SMA. The new items include activities such as shoulder abduction and flexion using weights that can be lifted by stronger patients. Other functional items included the ability to bring hands from lap to table and hand dexterity and manipulation items (picking

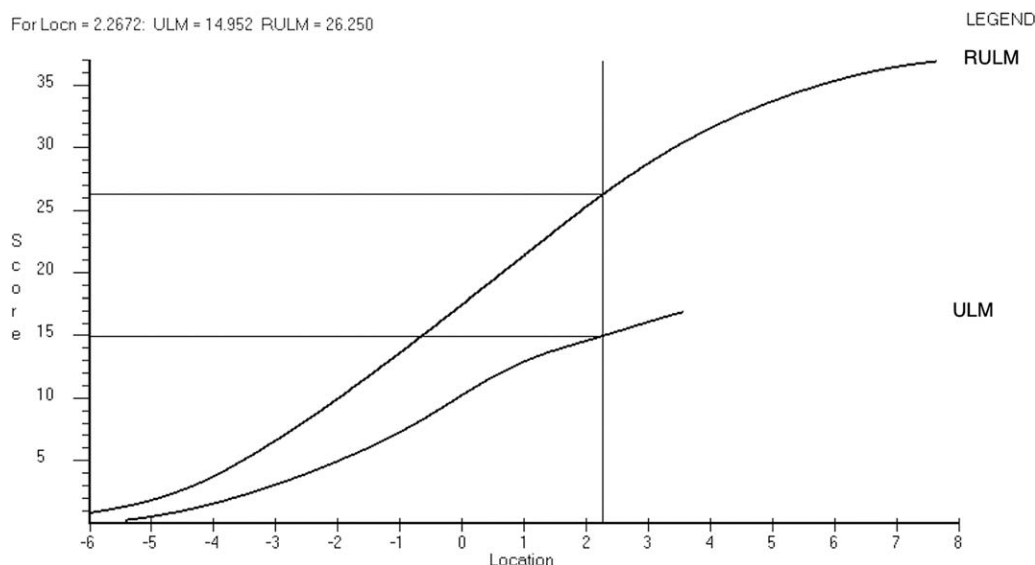


FIGURE 3. Equating of the 2 scales: RULM and ULM. The vertical axis denotes the raw score, thus a raw score on the ULM of 15 is equivalent to a raw score of 26 on the RULM. Entry item excluded from total score on RULM.

up coins and tearing a piece of paper) that further explore limited distal upper limb function.

The choice of items reflects what experienced clinicians and patients believed were clinically relevant. Patients with SMA and their caregivers were involved iteratively throughout the process to establish clinical meaningfulness and relevance of individual RULM items to activities of daily living. To ascertain whether the activities in the test were developmentally appropriate, they were piloted in the original version of the ULM in a cohort of typically developing children and could be successfully completed in all children, including the youngest ones, assessed at 30 months.⁵

The use of repeated Rasch analysis helped to improve the new scale robustness by highlighting possible issues that impact scale validity such as item fit, disordered thresholds or item redundancy. These issues were vetted by an expert panel of clinicians, and changes were made regarding the clinical relevance of the individual items and their scoring options. Plans for the future include collecting prospective data in a larger sample with the intent to undertake longitudinal analysis to review change scores over time as well as create a linearization of the RULM that would better ensure change across the breadth of the scale to mean the same regardless of patient ability. It is also possible to establish raw scores on the ULM that relate to RULM raw scores by means of a logit transformation as reported in Figure 3. This is a dynamic figure and is updated as more data are included to ensure greater precision and prediction.

We also plan to assess the impact of age, contractures, and gender on individual scores on this scale once a larger sample size is available.

Our preliminary findings, obtained as part of the training in the 3 networks, also suggest excellent inter-observer reliability. Ongoing studies are also in progress to establish different aspects of validation and cross correlation with other measures (myometry and grip) that are particularly relevant in weaker subjects. The RULM also is currently being used by all the participating centers as part of their natural history studies, which will allow better understanding of its sensitivity to detect change over time.

There was consensus among the clinical evaluators that the module is easy to perform and score, as categories are consistent. It was tolerated by participants, even younger children, allowing us to capture upper limb activities across a wide spectrum of both ambulatory and nonambulatory individuals with SMA across a range of functional

abilities. Analysis of the data clearly indicated that the scale proved to be a more suitable test for individuals with SMA and suitable for use in a multicenter setting because it fulfilled many key requirements.

The RULM can continue to identify change in motor function after an ambulatory individual becomes nonambulatory. These results together with the statistical robustness suggest that the RULM is a valuable addition to the existing scales that are currently being used in clinical trials in SMA.

Professor Giulio De Belvis, Statistician at the Catholic University of Rome, is acknowledged for his support for statistical analysis.

REFERENCES

- Bertini E, Burghes A, Bushby K, Estournet-Mathiaud B, Finkel RS, Hughes RA, et al. 134th ENMC International Workshop: outcome measures and treatment of spinal muscular atrophy, 11-13 February 2005, Naarden, The Netherlands. *Neuromuscul Disord* 2005;15:802-816.
- Mercuri E, Mayhew A, Muntoni F, Messina S, Straub V, Van Ommen GJ, et al. 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 2007, Naarden, The Netherlands; conjoint Institute of Myology TREAT-NMD meeting on physical activity monitoring in neuromuscular disorders, 11th July 2007, Paris, France. *Neuromuscul Disord* 2008;18:894-903.
- Kayadjanian N, Burghes A, Finkel RS, Mercuri E, Rouault F, Schwensen I, et al. SMA-EUROPE workshop report: opportunities and challenges in developing clinical trials for spinal muscular atrophy in Europe. *Orphanet J Rare Dis* 2013;8:44.
- Finkel R, Bertini E, Muntoni F, Mercuri E, Group ESWS. 209th ENMC International Workshop: outcome measures and clinical trial readiness in spinal muscular atrophy 7-9 November 2014, Heemskerk, The Netherlands. *Neuromuscul Disord* 2015;25:593-602.
- Mazzone E, Bianco F, Martinelli D, Glanzman AM, Messina S, De Sanctis R, et al. Assessing upper limb function in nonambulant SMA patients: development of a new module. *Neuromuscul Disord* 2011;21:406-412.
- Sivo S, Mazzone E, Antonaci L, De Sanctis R, Fanelli L, Palermo C, et al. Upper limb module in non-ambulant patients with spinal muscular atrophy: 12 month changes. *Neuromuscul Disord* 2015;25:212-215.
- Montes J, Glanzman AM, Mazzone ES, Martens WB, Dunaway S, Pasternak A, et al. Spinal muscular atrophy functional composite score: a functional measure in spinal muscular atrophy. *Muscle Nerve* 2015;52:942-947.
- Cano SJ, Mayhew A, Glanzman AM, Krossschell KJ, Swoboda KJ, Main M, et al. Rasch analysis of clinical outcome measures in spinal muscular atrophy. *Muscle Nerve* 2014;49:422-430.
- Hobart J, Cano S, Zajicek J, Thompson A. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol* 2007;6:1094-1105.
- Cano S, Warner T, Thompson A, Bhatia K, Fitzpatrick R, Hobart JC. The cervical dystonia impact profile (CDIP-58): can a Rasch developed patient reported outcome measure satisfy traditional psychometric criteria? *Health Qual Life Outcomes* 2008;6:58.
- Hobart J, Cano S. Improving the evaluation of therapeutic intervention in MS: the role of new psychometric methods. Monograph for the UK Health Technology Assessment Programme 2009;13:1-200.
- Mazzone ES, Vasco G, Palermo C, Bianco F, Galluccio C, Ricotti V, et al. A critical review of functional assessment tools for upper limbs in Duchenne muscular dystrophy. *Dev Med Child Neurol* 2012;54:879-885.
- Mayhew A, Mazzone ES, Eagle M, Duong T, Ash M, Decostre V, et al. Development of the Performance of the Upper Limb module for Duchenne muscular dystrophy. *Dev Med Child Neurol* 2013;55:1038-1045.
- Mercuri E, Mazzone ES. Choosing the right clinical outcome measure: from the patient to the statistician and back. *Neuromuscul Disord* 2011;21:16-19.