CLINICAL RESEARCH ARTICLE

MUSCLE&NERVE WILEY

Different trajectories in upper limb and gross motor function in spinal muscular atrophy

Giorgia Coratti PT, MsC[1,2] | Maria Carmela Pera MD, PhD[1,2] | Jacqueline Montes PT, EdD[3] | Amy Pasternak PT[4] | Mariacristina Scoto MD,PhD[5] | Giovanni Baranello MD, PhD[5] | Sonia Messina MD, PhD[6] | Sally Dunaway Young PT[7] | Allan M. Glanzman PT[8] | Tina Duong PT, PhD[7] | Roberto De Sanctis PT[2] | Elena Stacy Mazzone PT[1] | Evelin Milev PT[5] | Annemarie Rohwer PT[5] | Matthew Civitello PT[9] | Marika Pane MD, PhD[1,2] | Laura Antonaci MD[1,2] | Anna Lia Frongia MD[1,2] | Maria Sframeli MD, PhD[6] | Gian Luca Vita MD, PhD[6] | Adele D'Amico MD, PhD[10] | Irene Mizzoni PT[10] | Emilio Albamonte MD[11] | Basil T. Darras MD[4] | Enrico Bertini MD, PhD[10] | Valeria A. Sansone MD, PhD[11] | Francesca Bovis STAT, PhD[12] | John Day MD, PhD[7] | Claudio Bruno MD, PhD[13] | Francesco Muntoni MD, PhD[5,14] | Darryl C. De Vivo MD[3] | Richard Finkel MD[9] | Eugenio Mercuri MD, PhD[1,2]

[1]Pediatric Neurology, Università Cattolica del Sacro Cuore, Rome, Italy

[2]Centro Clinico Nemo, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy

[3]Departments of Rehabilitation and Regenerative Medicine and Neurology, Columbia University Irving Medical Center, New York, New York, USA

[4]Department of Neurology, Boston Children's Hospital, Harvard Medical School, Boston, Massachusetts, United States

[5]Dubowitz Neuromuscular Centre, UCL Institute of Child Health & Great Ormond Street Hospital, London, UK

[6]Department of Clinical and Experimental Medicine and Centro Clinico Nemo Sud, University of Messina, Messina, Italy

[7]Department of Neurology, Stanford University, Stanford, California, USA

[8]Department of Physical Therapy, Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA

[9]St. Jude Children's Research Hospital, Memphis, Tennessee, USA

[10]Unit of Neuromuscular and Neurodegenerative Disorders, Department of Neurosciences, IRCCS Bambino Gesù Children's Hospital, Rome, Italy

[11]Neurorehabilitation Unit, University of Milan, Neuromuscular Omnicentre Clinical Center, Niguarda Hospital, Milan, Italy

[12]Biostatistics Unit, Department of Health Sciences, University of Genoa, Genoa, Italy

[13]Center of Experimental and Translational Myology, IRCCS Istituto Giannina Gaslini, Genoa, Italy

[14]NIHR Great Ormond Street Hospital Biomedical Research Centre, London, UK

Correspondence
Eugenio Mercuri, Pediatric Neurology, Catholic University, Largo Gemelli 8, 00168 Rome, Italy.
Email: eugeniomaria.mercuri@unicatt.it

Abstract

Introduction: The Hammersmith Functional Motor Scale Expanded (HFMSE) and the Revised Upper Limb Module (RULM) have been widely used in natural history studies and clinical trials. Our aim was to establish how the scales relate to each other at different age points in spinal muscular atrophy (SMA) type 2 and 3, and to describe their coherence over 12 mo.

Methods: The study was performed by cross-sectional and longitudinal reanalysis of previously published natural history data. The longitudinal analysis of the 12-mo changes also included the analysis of concordance between scales with changes grouped as stable (±2 points), improved (>+2) or declined (>−2).

Results: Three hundred sixty-four patients were included in the cross-sectional analysis, showing different trends in score and point of slope change for the two scales. For type 2, the point of slope change was 4.1 y for the HFMSE and 5.8 for the RULM, while for type 3, it was 6 y for the HFMSE and 7.3 for the RULM. One-hundred-twenty-one patients had at least two assessments at 12 mo. Full concordance was found in 57.3% of the assessments, and in 40.4% one scale remained stable and the other changed. Each scale appeared to be more sensitive to specific age or functional subgroups.

Discussion: The two scales, when used in combination, may increase the sensitivity to detect clinically meaningful changes in motor function in patients with SMA types 2 and 3.

KEYWORDS
disease severity, motor, neuromuscular disorders, outcome measures, spinal muscular atrophy

1 | INTRODUCTION

Spinal muscular atrophy (SMA) is caused by mutations in the survival motor neuron 1 gene (*SMN1*) on chromosome 5q.[1] Historically, SMA has been classified into different types according to onset and maximum motor function achieved, with types 1 to 3 being the most frequent forms with pediatric onset.[2] The advent of clinical trials has highlighted the need to have reliable tools to measure functional changes in SMA patients.[3] Two disease specific scales, in their original forms or their revised versions, have been used in most trials in patients with later-onset SMA. The Hammersmith Functional Motor Scale,[4] and the expanded version, HFMSE,[5] have been used to assess gross motor function. Upper limb function has been assessed using the Upper Limb Module (ULM)[6] and more recently, its revised version, the Revised Upper Limb Module (RULM).[7] We and others have reported longitudinal data using the HFMSE in both types 2 and 3, exploring different statistical analyses to establish the patterns of progression in large cohorts as part of international efforts.[8-15] Less has been reported on the RULM also because the scale was developed more recently.

Although both scales have been used in several natural history studies[8-15] and in clinical trials, the RULM and the HFMSE results have always been analyzed separately,[16] and their relationship to one another has not been systematically explored.

Only one study explored the correlation between the HFMSE and upper limb function but used the previous version of the module (ULM).[17] In that study, we reported a good correlation (Pearson: 0.65) between the HFMSE and the ULM, highlighting how the ULM appeared to fill the gap caused by the floor effect observed in the HFMSE in patients who had lost the ability to sit or were just able to sit for a few seconds. In contrast, the ULM appeared to have a ceiling effect in stronger patients, as it had been specifically designed for use in young weak children.[17] The RULM was developed to overcome some of these shortcomings[7] and as the scales are being increasingly used in clinical trials, it has become important to assess how the two scales relate to each other. In this study we analyzed the results obtained in a large international cohort of type 2 and 3 patients assessed with both HFMSE and RULM in order to establish their correlations with one another. We also analyzed the 12-mo changes on the two scales and the level of concordance between them, in an attempt to identify possible patterns of discordance in relation to age, functional status, or SMA type.

2 | METHODS

The study was performed using cross-sectional and longitudinal data collected by three SMA networks in the United States, Italy, and UK from September, 2015, to January, 2020.[18]

The study was approved by the Ethical Committee of each center (*Catholic University of Sacred Heart, Fondazione Policlinico Universitario Agostino Gemelli IRCCS; Columbia University; Boston Children's Hospital; UCL Institute of Child Health & Great Ormond Street Hospital; University of Messina; Stanford University; Children's Hospital of Philadelphia; IRCCS Bambino Gesù Children's; University of Milan, Niguarda Hospital; IRCCS Istituto Giannina Gaslini*). Parents of participants (of minors) and/or patients were informed that the data collected as part of our routine clinical assessment were going to be used anonymously for an observational study defining the natural history of the diseases and all provided informed consent/assent.

All patients had a genetically confirmed diagnosis of SMA, were not treated with experimental or approved drugs (nusinersen, risdiplam, onasemnogene abeparvovec-xioi) and only those with a diagnosis of type 2 and 3 SMA were included in the study. Patients received best available supportive care, as outlined in the SMA standard-of-care consensus statement,[19,20] that was generally applied in a uniform way among the sites, depending partly upon local healthcare practices. Patients who had corrective surgery for scoliosis during the period of observation were excluded from this analysis.

2.1 | HFMSE

The scale consists of 33 items, investigating the patient's ability to perform various activities.[5] Each activity, categorized in items, is scored on a 3-point system, with a score of 2 for "performs without modification", 1 "performs with modification/adaption/compensation" and 0 for "unable to perform the task". Total score is the sum of the item's individual scores and can range from 0, if all the activities are failed, to 66, if all the activities are completed. All items were tested without spinal jacket or orthoses.

2.2 | RULM

The scale includes an entry item to establish functional levels and 19 items covering distal to proximal upper limb function.[7] Eighteen of these 19 items are scored using a 3-point system and 1 item is scored using a 2-point system. The total score ranges from 0, if no activities can be completed, to 37, if all the activities are achieved fully without any compensation or task modification.

2.3 | Evaluator training sessions

The physical therapists in the participating clinics received the same training programs with establishment of yearly intra- and inter-rater reliability, and standardized procedures of scale administration. Inter- and intra observer reliability of the HFMSE and RULM have already been reported.[21]

2.4 | Statistical analysis

The relationship between HFMSE and RULM was evaluated cross-sectionally, including all patients and all assessments. Despite the limitations of dichotomizing SMA into two types (types 2 and 3), as this is rather a spectrum of disease with respect to age of onset and severity, we analyzed the results in the two types separately in order to allow comparison with previous studies.[8,9,12,16]

Summary statistics (N, mean, SD, range) were used. The cohort was subdivided into age subgroups on the basis of cutoff points identified on our previous observations reporting slopes of progression in types 2 and 3 at different age points.[10,22] The same criteria were used to group patients by SMA type and baseline functional status (non-sitter, sitter, walkers).[10,22]

A polynomial line was assessed in order to graphically describe the population trends on the cross-sectional dataset; age of slope change was found by calculating the local maxima of the curve from the curve equation. Spearman' correlations were assessed between scales.

In order to describe the HFMSE/RULM trends over 12-mo, as performed for other papers in patients with long-term follow-up,[10,12] we included multiple 12-mo paired assessments from individual patients. Each interval was considered independently, with a new age point and HFMSE/RULM baseline for each calculated interval. Mann–Whitney *U* test was performed in order to compare 12 mo changes on HFMSE and RULM between SMA type 2 and 3.

We also analyzed the data with a different approach. As previously reported in other papers[9,12] 12-mo changes were clustered in three groups: patients with stable results (± 2 points), those with losses of more than 2 points and those with improvements of more than 2 points. We examined if the scores on each scale were stable (±2 points) improved (>+2) or declined (>−2).

The percentages of patients within each group were compared across age classes, functional status, and SMA type. Tables of contingency were used in order to describe paired assessments that showed fully concordant changes, that is, stable on both HFMSE and RULM; fully discordant changes, that is, improving on the HFMSE while declining on the RULM or improving on the RULM while declining on the HFMSE; and the ones in which one scale was stable while the other was moving. Chi-squared testing was used to analyze the distribution by SMA type of fully concordant changes, fully discordant changes or the ones in which only one scale changed, while Cohen kappa was used to assess the agreement of distribution of fully concordant changes, fully discordant changes or the in which only one scale changed.

3 | RESULTS

3.1 | Cross-sectional cohort

Eight hundred thirty-seven assessments from 364 patients were included in the study. Baseline characteristics of the enrolled participants are presented in Table 1. Of the 837 assessments, 462 were

TABLE 1 Baseline characteristics of the enrolled participants

		TYPE 2 (N:213)			TYPE 3 (N:151)			
		All (N: 213)	Non-sitter (N:47)	Sitter (N:166)	All (N: 151)	Non-sitter (N:4)	Sitter (N:64)	Walker (N:83)
Age (y)	MEAN (SD)	11.12 (+6.82)	16.85 (+7.18)	9.5 (+5.78)	12.95 (+6.8)	17.26 (+7.03)	15.44 (+6.71)	10.82 (+6.16)
	MIN; MAX	2.5; 29.61	4; 29.61	2.5; 27.7	2.5; 28.76	10.6; 23.4	2.5; 28.73	2.88; 28.76
Gender	Male (n)	117	27	90	73	3	34	36
	Female (n)	96	20	76	78	1	30	447
SMN2 copy number	1 (n)	1	1	0	1	0	0	1
	2 (n)	25	9	16	7	0	2	5
	3 (n)	111	17	94	51	1	25	25
	≥4 (n)	4	0	4	29	0	10	19
	Unknown (n)	72	20	52	63	3	27	33
HFMSE at 1st visit	MEAN (SD)	8.92 (8.32)	0.45 (1.00)	11.32 (7.91)	39.43 (17.89)	1.5 (0.58)	25.09 (12.77)	52.31 (7.83)
	MIN; MAX	0;36	0;4	2;36	1;66	1;2	2;53	31;66
RULM at 1st visit	MEAN (SD)	13.67 (7.22)	4.94 (5.43)	16.14 (5.56)	29.93 (7.03)	11 (5.89)	26.20 (6.29)	33.72 (3.99)
	MIN; MAX	0;30	0;18	4;30	3;37	3;17	13;37	22;37

available from 213 type 2 patients and 375 were available from 151 type 3 patients. All the patients included in the present study have already been reported in previous studies reporting one of the two scales.[8-13]

3.1.1 | SMA 2

In the type 2 patients, 367 of 462 assessments (79.4%) were from sitters and 95 (20.5%) from non-sitters (patients who had lost the ability to sit). No patients had spinal surgery prior to age 5 y. The HFMSE point of slope change was at 4.1 y and the RULM point of slope change was 5.8 y (Figure 1A). The correlation between HFMSE and RULM in the overall type 2 assessments was 0.730. Supporting Information Table S1, which is available online, reports details on the mean age, HFMSE and RULM of the type 2 cohort according to age groups and functional status (non-sitter, sitter). Table 2 reports correlation between the scales by functional status.

3.1.2 | SMA 3

In the type 3 patients, 218 of 375-assessments (58.1%) were from walkers, 152 (40.5%) from sitters (who lost the ability to walk), and 5 (0.01%) from non-sitters. No type 3 participants had spinal surgery prior to age 7. The HFMSE point of slope change was 6 y and the RULM point of slope change was 7.3 y (Figure 1B). The correlation between HFMSE and RULM was 0.787. Supporting Information Table S2 reports details on the mean age, HFMSE and RULM of the type 3 cohort according to age groups and functional status (sitter, walker). Table 2 reports correlations between the scales by functional status.

3.2 | Longitudinal observations: 12-mo changes and correlation between scales

One hundred twenty-one patients with ages ranging from 2.67 to 29.61 y (mean 11.76; SD 6.07) had at least two assessments at 12-mo interval, with 63 having more than one paired assessment (mean:1.89, SD: 1.10), for a total of 225 12-mo paired assessments.

The correlation between HFMSE and RULM 12-mo changes in the whole cohort was 0.214 (0.186 and 0.253 in type 2 and type 3, respectively). No difference was found in 12-mo mean change between SMA 2 and SMA 3 for HFMSE ($P = .820$) and RULM ($P = .159$).

3.2.1 | SMA 2

One-hundred twelve assessments were available from 65 type 2 patients. Three patients classified as "sitters" at baseline, lost their ability to sit during the 12-mo follow-up.
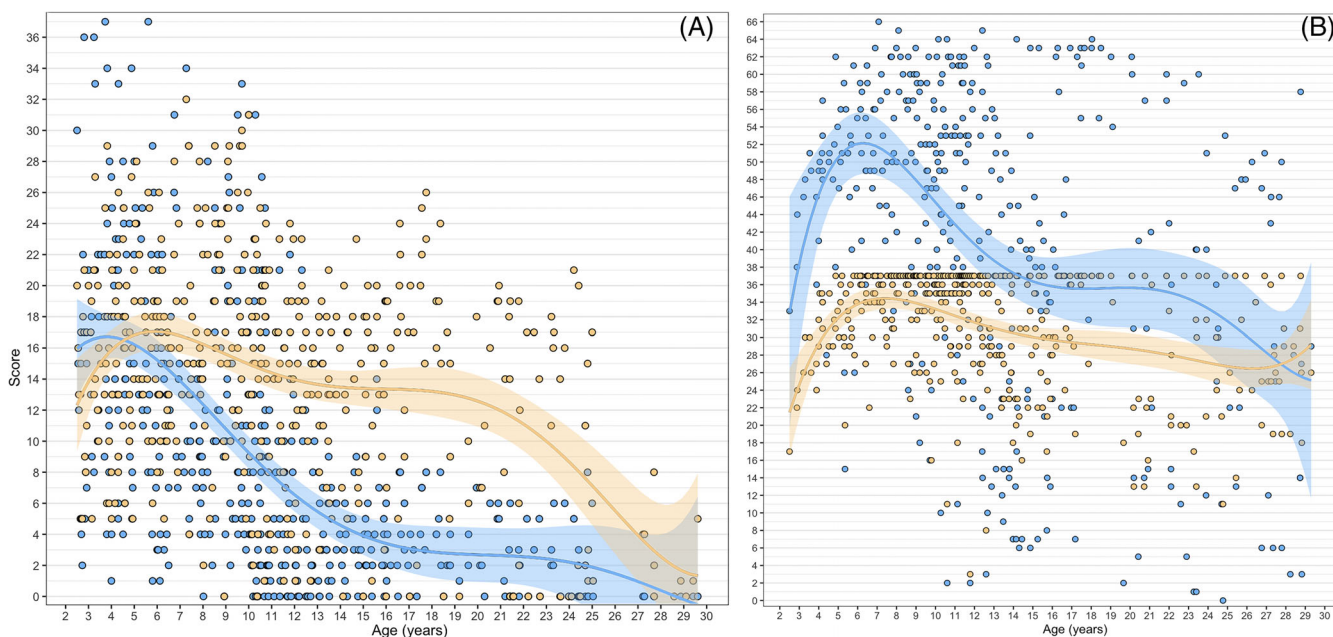
FIGURE 1 Distribution of scores according to age and total score for SMA 2 patients (panel A) and SMA 3 patients (panel B). Color coding = blue: HFMSE, orange: RULM. Blue line: polynomial line for HFMSE (ribbon: 95% confidence interval [CI]). Orange line: polynomial line for RULM (ribbon: 95% CI). Polynomial line describes progression overtime, local maxima of the curve is indicative of the point of slope

TABLE 2 Correlations between HFMSE and RULM by SMA type and functional status

		Spearman rho
SMA 2	All	0.730
	Non-sitter	0.422
	Sitter	0.629
SMA 3	All	0.787
	Non-sitter	0.548
	Sitter	0.800
	Walker	0.493

For the cohort, the HFMSE 12-mo changes was −0.54 (SD 2.75) and the RULM change was −0.13 (SD 3.12). Supporting Information Table S3 reports details of descriptive statistics on 12-mo changes of the SMA 2 cohort subdivided by age group and functional status (non-sitter, sitter).

3.2.2 | SMA 3

One-hundred thirteen assessments were available from 56 type 3 patients. One patient classified as a "walker" at baseline lost ambulation during the 12-mo follow-up. For the cohort, the HFMSE 12 mo was −0.85 (SD 3.93) and the RULM change was 0.29 (SD 2.47). Supporting Information Table S4 reports details of descriptive statistics on 12-mo changes for the SMA 3 cohort subdivided by age group and functional status (sitter, walker).

3.3 | Concordance between HFMSE and RULM changes

The 12-mo changes were also computed in a contingency table to verify percentage of concordance and discordance regarding decline, stability, and improvements between the HFMSE and RULM. The distribution of stable, declined, or improved assessment on the HFMSE was different by SMA type ($X2$ [2, $N = 264$] = 17.294, <0.001) with type 2 being overall more stable than type 3 (75.96% vs. 51.84%, respectively). This held true even when patients with scores indicative of a floor score (scores lower than 2) ($X2$ [2, $N = 225$] = 7.260, $P = .027$), or ceiling score (score = 66) ($X2$ [2, $N = 264$] = 17.294, $P = <.001$) at baseline were excluded. No difference in distribution was observed on the RULM by SMA type ($X2$ [2, $N = 264$] = 4.773, $P = .092$), even when excluding the ones with floor score (scores lower than 2) ($X2$ [2, $N = 255$] = 5.874, $P = .054$) or ceiling score (score = 37) ($X2$ [2, $N = 220$] = 1.373, $P = .503$) at baseline.

Table 3 reports the table of contingency for HFMSE and RULM subdivided by SMA type. More details on results by age groups can be found in Supporting Information Table S5.

In type 2, concordance between the RULM and the HFMSE 12-mo change (both declining, stable, or improving) was found in 67/112 (59.82%) paired 12-mo assessments. Fully discordant results (one scale improving and the other one declining) were found in 3/112 (2.67%), with 2 of the 3 showing improved RULM while declining on the HFMSE ($\kappa = 0.089$, $P = .187$).

In 42 patients (37.50%), one of the two scales remained stable while the other had changes >2 points. In 24 of the 42, the HFMSE

TABLE 3 SMA II and III table of contingency for HFMSE and RULM subdivided by SMA type

		RULM		
SMA type	HFMSE	Decline	Stability	Improvement
SMA 2 (N:112)	Decline	4 (3.57%)	13 (11.60%)	2(1.78%)
	Stability	14 (12.50%)	61 (54.46%)	10 (8.92%)
	Improvement	1 (0.89%)	5 (4.46%)	2 (1.78%)
SMA 3 (N:113)	Decline	4 (3.54%)	28 (24.77%)	2 (1.76%)
	Stability	4 (3.54%)	52 (46.01%)	3 (3.54%)
	Improvement	0 (0.88%)	14 (12.38%)	6 (5.31%)

was stable, with the RULM decreasing in 14/24 and increasing in 10. In the 24 with stable HFMSE, 6 had a floor score (HFMSE score <2) at baseline, with the RULM decreasing in 5/6 and increasing in 1. None reached a ceiling score (HFMSE score = 66).

In 18/42, the RULM was stable, with the HFMSE decreasing in 13/18 and increasing in 5. In the 18 with stable RULM, none reached a floor (RULM score <2) or ceiling (RULM score = 37) score.

In type 3 patients, concordance between the RULM and the HFMSE 12-mo change (both declining, stable, or improving) was found in 59/113 (52.21%) paired 12-mo assessments. Fully discordant results (one scale improving and the other one declining) were found in 2/113 (1.76%) showing improvements on the RULM, while declining on the HFMSE ($\kappa = 0.179$, $P = .001$).

In 49 patients (43.36%), one of the 2 scales remained stable, while the other had changes >2 points. In 7/49, the HFMSE was stable, with the RULM decreasing in 4/7 and increasing in 3.

In the seven with stable HFMSE, none had floor (HFMSE score <2) or ceiling (HFMSE score = 66) scores.

In 42/49, the RULM was stable, with the HFMSE decreasing in 28/42 and increasing in 14. In the 42 with stable RULM, none had a floor score (HFMSE score <2), while 11 had the maximum RULM score (37 points).

4 | DISCUSSION

The results of our cross-sectional analysis confirm the overall correlation between the HFMSE and upper limb function previously observed using the original version of the upper limb module in a smaller cohort of 74 non-ambulant patients.[8] In this paper, the use of the RULM in a larger cohort, also including ambulant patients, allowed us more detailed information, overcoming some of the limits of the ULM, a scale that was not designed for stronger ambulant patients.

Our results clearly indicate that the scores on the two scales show different trends over time. In both type 2 and type 3 patients, the peak indicating the highest scores achieved prior the subsequent decline occurred at a younger age on the HFMSE compared with the RULM. The mean decline in RULM scores was overall milder.

When we analyzed the 12 mo changes, we also found that the changes in the two scales often occurred at different times in the two SMA types, and the two scales were complementary to detect changes in one or the other domain. This was particularly true in

subgroups in whom one of the two scales was relatively stable because of floor or ceiling effects. After the age of 13 y in type 2 patients the HFMSE scores are generally low and relatively stable, and the changes are mainly observed on the RULM. In contrast, many type 3 patients reach top scores on the RULM, that often remain stable, and the changes are mainly observed on the HFMSE. This may partly explain the low correlation of the 12 mo changes in the two scales that is probably also partly due to the fact that, even when trending in the same direction, the magnitude of the changes was often different. In order to establish the overall level of concordance between the changes on the two scales, we also used a different approach. Changes beyond ±2 points on both HFMSE and RULM have previously been reported as clinically meaningful.[8,9,12,16,23] This has also been confirmed by studies reporting patient perspective. We therefore also analyzed the level of concordance between the two scales by using these criteria. In both type 2 and 3 full concordance was found in approximately half of the assessments while full discordance was rare. It is of note that, in type 2 SMA below the age of 7 y, we never observed a concurrent deterioration on both scales; conversely, after the age of 5 y, with the exception of one patient, there was never an improvement on both scales.

In the great majority of the remaining cases, one scale remained stable while the other changed. This relative discordance largely reflects the differences observed on the curves reporting the scores at different age in the two scales.

These results indicate that the two scales are in some way complementary to each other, and when used in combination, not only capture relevant domains of function of these patients, but may also increase the possibility to detect changes at different ages and functional levels. The changes occurring at different times in the two scales reflect the different gradient of weakness and progression between upper and lower limbs, with the lower limb being often more involved at earlier stages of the disease. These findings are also in keeping with pathological findings reporting that the motor neurons in the anteromedial zone at the cervical segments of the spinal cord are relatively more preserved than those in the other segments.[24]

This raises the question on how to use the two scales in combination to assess possible changes following interventions in large cohorts including patients at different ages and with different SMA types and functional levels. One option could be that of counting responders on each scale and in combination. The alternative could be to develop different scoring systems, combining the two different

domains as previously used in the Motor Function Measure-32 (MFM 32),[25,26] a scale that covers a large spectrum of activity that was, however, not specifically designed for SMA. For the stronger patients, other measures, such as the 6 Minute Walk Test (6MWT), may further contribute to the assessment of possible changes. This has been also suggested in the past by a paper that combined the HFMSE, ULM, and 6MWT.[27] In that publication, statistical methods similar to those used in the development of the MFM[25,28,29] were used to assess the relative contributions of the individual scales to the total composite score.

Combining the two scales would help identify patients in whom the possibility of showing a response over time may be limited in one of the two scales. Combining the scales may, more generally, be used for patient stratification or for trial design in future interventional studies.

Work is in progress to use appropriate statistical approaches to combine the existing scales, in order to increase the possibility of detecting any possible change in function in types 2 and 3 SMA patients of different ages and functional levels.

ETHICAL PUBLICATION STATEMENT

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sonia Messina https://orcid.org/0000-0001-7994-3391

REFERENCES

1. D'Amico A, Mercuri E, Tiziano FD, Bertini E. Spinal muscular atrophy. *Orphanet J Rare Dis*. 2011;6:71.
2. Dubowitz V. Chaos in classification of the spinal muscular atrophies of childhood. *Neuromuscul Disord*. 1991;1(2):77-80.
3. 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(7):593-602.
4. Main M, Kairon H, Mercuri E, Muntoni F. The Hammersmith functional motor scale for children with spinal muscular atrophy: a scale to test ability and monitor progress in children with limited ambulation. *Eur J Paediatr Neurol*. 2003;7(4):155-159.
5. O'Hagen JM, Glanzman AM, McDermott MP, et al. An expanded version of the Hammersmith Functional Motor Scale for SMA II and III patients. *Neuromuscul Disord*. 2007;17(9–10):693-697.
6. Mazzone E, Bianco F, Martinelli D, et al. Assessing upper limb function in nonambulant SMA patients: development of a new module. *Neuromuscul Disord*. 2011;21(6):406-412.
7. Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: development of a new module. *Muscle Nerve*. 2017;55(6):869-874.
8. Pera MC, Coratti G, Mazzone ES, et al. Revised upper limb module for spinal muscular atrophy: 12 month changes. *Muscle Nerve*. 2019;59(4):426-430.
9. Coratti G, Lucibello S, Pera MC, et al. Gain and loss of abilities in type II SMA: a 12-month natural history study. *Neuromuscul Disord*. 2020;30(9):765-771.
10. Coratti G, Pera MC, Lucibello S, et al. Age and baseline values predict 12 and 24-month functional changes in type 2 SMA. *Neuromuscul Disord*. 2020;30(9):756-764.
11. Mazzone ES, Mayhew A, Montes J, et al. Revised upper limb module for spinal muscular atrophy: development of a new module. *Muscle Nerve*. 2016;55(6):869-874.
12. Mercuri E, Finkel R, Montes J, et al. Patterns of disease progression in type 2 and 3 SMA: implications for clinical trials. *Neuromuscul Disord*. 2016;26(2):126-131.
13. Mercuri E, Lucibello S, Pera MC, et al. Long-term progression in type II spinal muscular atrophy: a retrospective observational study. *Neurology*. 2019;93(13):e1241-e1247.
14. Kaufmann P, McDermott MP, Darras BT, et al. Observational study of spinal muscular atrophy type 2 and 3: functional outcomes over 1 year. *Arch Neurol*. 2011;68(6):779-786.

15. Kaufmann P, McDermott MP, Darras BT, et al. Prospective cohort study of spinal muscular atrophy types 2 and 3. *Neurology*. 2012;79 (18):1889-1897.

16. Mercuri E, Darras BT, Chiriboga CA, et al. Nusinersen versus sham control in later-onset spinal muscular atrophy. *N Engl J Med*. 2018; 378(7):625-635.

17. Sivo S, Mazzone E, Antonaci L, et al. Upper limb module in non-ambulant patients with spinal muscular atrophy: 12 month changes. *Neuromuscul Disord*. 2015;25(3):212-215.

18. Mercuri E, Finkel R, Scoto M, et al. Development of an academic disease registry for spinal muscular atrophy. *Neuromuscul Disord*. 2019; 29(10):794-799.

19. Finkel RS, Mercuri E, Meyer OH, et al. Diagnosis and management of spinal muscular atrophy: part 2: pulmonary and acute care; medications, supplements and immunizations; other organ systems; and ethics. *Neuromuscul Disord*. 2017;28(3):197-207.

20. Mercuri E, Finkel RS, Muntoni F, et al. Diagnosis and management of spinal muscular atrophy: part 1: recommendations for diagnosis, rehabilitation, orthopedic and nutritional care. *Neuromuscul Disord*. 2017; 28(2):103-115.

21. Glanzman AM, Mazzone ES, Young SD, et al. Evaluator training and reliability for SMA global Nusinersen trials. *J Neuromuscul Dis*. 2018;5 (2):159-166.

22. Coratti G, Messina S, Lucibello S, et al. Clinical variability in spinal muscular atrophy type III. *Ann Neurol*. 2020;88(6):1109-1117.

23. Pera MC, Coratti G, Forcina N, et al. Content validity and clinical meaningfulness of the HFMSE in spinal muscular atrophy. *BMC Neurol*. 2017;17(1):39.

24. Araki S, Hayashi M, Tamagawa K, et al. Neuropathological analysis in spinal muscular atrophy type II. *Acta Neuropathol*. 2003;106(5):441-448.

25. Berard C, Fermanian J, Payan C. Outcome measure for SMA II and III patients. *Neuromuscul Disord*. 2008;18(7):593-594; author reply 594–595.

26. Berard C, Payan C, Hodgkinson I, et al. A motor function measure for neuromuscular diseases. Construction and validation study. *Neuromuscul Disord*. 2005;15(7):463-470.

27. Montes J, Glanzman AM, Mazzone ES, et al. Spinal muscular atrophy functional composite score: a functional measure in spinal muscular atrophy. *Muscle Nerve*. 2015;52(6):942-947.

28. Berard C, Payan C, Fermanian J, Girardot F, Groupe d'Etude MFM. A motor function measurement scale for neuromuscular diseases - description and validation study. *Rev Neurol (Paris)*. 2006;162(4): 485-493.

29. Vuillerot C, Payan C, Iwaz J, Ecochard R, Berard C, Group MFMSMAS. Responsiveness of the motor function measure in patients with spinal muscular atrophy. *Arch Phys Med Rehabil*. 2013;94(8):1555-1561.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.