



Published in final edited form as:

*Neuromuscul Disord.* 2015 July ; 25(7): 554–562. doi:10.1016/j.nmd.2015.03.008.

## A functional scale for spinal and bulbar muscular atrophy: Cross-sectional and longitudinal study

Atsushi Hashizume<sup>a</sup>, Masahisa Katsuno<sup>a</sup>, Keisuke Suzuki<sup>a</sup>, Haruhiko Banno<sup>a,b</sup>, Noriaki Suga<sup>a</sup>, Tomoo Mano<sup>a</sup>, Amane Araki<sup>a</sup>, Yasuhiro Hijikata<sup>a</sup>, Christopher Grunseich<sup>c</sup>, Angela Kokkinis<sup>c</sup>, Akihiro Hirakawa<sup>d</sup>, Hirohisa Watanabe<sup>a</sup>, Masahiko Yamamoto<sup>e</sup>, Kenneth H. Fischbeck<sup>c</sup>, and Gen Sobue<sup>a,\*</sup>

<sup>a</sup>Department of Neurology, Nagoya University Graduate School of Medicine, Nagoya 466-8550, Japan

<sup>b</sup>Institute for Advanced Research, Nagoya University, Nagoya 464-8601, Japan

<sup>c</sup>Neurogenetics Branch, National Institute of Neurological Disorders and Stroke, National Institutes of Health, Bethesda, MD 20892, USA

<sup>d</sup>Biostatistics Section, Center for Advanced Medicine and Clinical Research, Nagoya University Graduate School of Medicine, Nagoya, Japan

<sup>e</sup>Department of Speech Pathology and Audiology, Aichi-Gakuin University School of Health Science, 12 Araiike, Iwasaki-cho, Nisshin 470-0195, Japan

### Abstract

We aimed to develop, validate, and evaluate a disease-specific outcome measure for SBMA: the Spinal and Bulbar Muscular Atrophy Functional Rating Scale (SBMAFRS). We examined the Japanese version (SBMAFRS-J) in 80 Japanese SBMA subjects to evaluate its validity and reliability. We then assessed this scale longitudinally in 41 additional SBMA subjects. The English version (SBMAFRS-E) was also tested in 15 US subjects. The total score of the SBMAFRS-J was distributed normally without an extreme ceiling or floor effect. For SBMAFRS-J, the high intra- and inter-rater agreement was confirmed (intra-class correlation coefficients [ICCs] 0.910 and 0.797, respectively), and internal consistency was satisfactory (Cronbach's alpha 0.700–0.822). In addition, SBMAFRS-J demonstrated concurrent, convergent, and discriminant validity, except for the respiratory subscale. The inter-rater reliability and internal consistency of SBMAFRS-E were also satisfactory. Longitudinally, SBMAFRS-J showed a higher sensitivity to disease progression than the existing clinical measures. In conclusion, we developed and validated a disease-specific functional rating scale for SBMA in both Japanese and English versions, although it needs to be re-assessed in interventional studies with a larger sample size including English speaking subjects.

\*Corresponding author. Department of Neurology, Nagoya University Graduate School of Medicine, 65 Tsurumai-cho, Showa-ku, Nagoya 466-8550, Japan. Tel.: +81 52 744 2385; fax: +81 52 744 2384. sobueg@med.nagoya-u.ac.jp (G. Sobue).

### Appendix: Supplementary material

Supplementary data to this article can be found online at [doi:10.1016/j.nmd.2015.03.008](https://doi.org/10.1016/j.nmd.2015.03.008).

## Keywords

Spinal and bulbar muscular atrophy; Validation; Functional rating scale; Motor neuron disease

---

## 1. Introduction

Spinal and bulbar muscular atrophy (SBMA), also known as Kennedy's disease, is an adult-onset, hereditary neuromuscular disease characterized by extremity muscle atrophy, weakness, contraction fasciculation, and bulbar involvement [1–3]. The progression of neurological deficits is usually slow in SBMA, with the average interval between the onset of symptoms and death approximately 20 years [4,5].

There is currently no effective treatment to slow the progression of SBMA. Although clinical trials of potential therapies have been done [6–9], definite efficacy has not been clearly demonstrated in randomized controlled trials. These results appear to be partly attributable to the absence of established outcome measures that are sensitive to the disease-specific symptoms as well as the limited number of patients, which may diminish statistical power. Namely, since the progression of SBMA is slow, existing outcome measures such as those for amyotrophic lateral sclerosis (ALS) are not sensitive enough to detect the deterioration of motor function in the patients [10].

Thus, the present study aimed to develop and validate a disease-specific functional rating scale for SBMA (SBMAFRS) and to evaluate whether this scale is more advantageous than other existing outcome measures. In addition, for future use in global clinical studies, the English version of SBMAFRS was also evaluated for reliability.

## 2. Methods

### 2.1. Development of the SBMAFRS

The SBMAFRS was designed as a quantitative outcome measure of global disability in SBMA, especially for use in clinical trials. According to this concept, we developed the SBMAFRS using the following four stages: item generation and scale development; cross-cultural translation; scale validation; and scale evaluation.

At the item generation and scale development stage, we first listed up all the items of the revised Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS-R) and those of the modified Norris Scale, and created novel items regarding bulbar and truncal function, which were not included in the existing scales. Then seven experts of SBMA selected as many candidate items as possible with consideration of the score distribution in our previous studies [7–10]. From the candidates, the best suitable fourteen items were selected and modified to fit the symptoms and severity of SBMA if necessary. Correspondence of the items of SBMAFRS and those of the existing scales was shown in Table SA-1. In the process of determining item balance, we gave weight to disease-specific features and their clinical relevance to SBMA. In the SBMAFRS, 5 of the 14 items addressed bulbar-related symptoms because these are the primary symptoms of SBMA and are occasionally lethal. Although pulmonary function tends to be preserved in SBMA, one breathing-related item

was adopted, because serious respiratory dysfunction has been reported in certain subsets of patients and associated with the prognosis [5,8]. This strategy may permit the identification of specific areas that respond to treatment or the sub-classification of SBMA into etiologically or therapeutically relevant subtypes.

We designed alternatives for each item so that the scale was sensitive enough to detect subtle neurological changes. For example, although moderate dysphagia that requires a change of dietary consistency is scored as 3 in the ALSFRS-R, we subdivided this status into the following 2 scores in the SBMAFRS: score 2, avoids some particular textures of food that may cause choking, etc.; and score 1, food textures always have to be modified to a soft or chopped consistency. In addition, we also ensured that the expressions used in the alternatives were as clear as possible to avoid ambiguity. For instance, a specified frequency was required to choose the alternative in the item for control of salivation: score 3, less than once a week; and score 2, not less than once a week.

As a result, a functional rating scale consisting of 14 items was generated. At the cross-cultural translation stage, we translated the original Japanese version (SBMAFRS-J) into English (SBMAFRS-E) for the common use of this scale in English-speaking countries. The translation procedure followed standard methods [11]. The SBMAFRS-J was translated into English considering cultural differences by a bilingual translator, and it was later submitted to an expert committee. This committee, consisting of C.G., G.S., and M.K., reviewed and revised the linguistic and conceptual equivalence of the original Japanese version and the first English version. Next, this English version was back translated into Japanese by a Japanese native without *a priori* knowledge of the intent and concept of this scale. Finally, the expert committee reviewed all reports and confirmed the linguistic and conceptual equivalence of this scale, and completed the English version. At the scale validation stage, we assessed the SBMAFRS-J on 80 subjects with SBMA, and evaluated its validity and reliability. We also assessed the reliability of the SBMAFRS-E on 15 subjects with SBMA in the USA. At the scale evaluation stage, we performed a 48-week longitudinal analysis on an additional 41 Japanese subjects with SBMA to evaluate the usefulness of the SBMAFRS-J as a potential endpoint in clinical trials. The evaluators of the SBMAFRS were physicians or research nurses who were familiar with the medical care for SBMA.

## 2.2. Subjects and validation procedure

We recruited a total of 80 consecutive male Japanese subjects with SBMA as the SBMAFRS-J validation group. The inclusion criteria were as follows: 1) genetically confirmed male Japanese SBMA subjects with more than one of the following findings: muscle weakness, muscle atrophy, and bulbar dysfunction; 2) subjects who were 25–75 years old at the time of informed consent; and 3) subjects who had not undergone surgical castration.

To assess inter-rater reliability, a total of 80 Japanese participants were assessed twice by two independent evaluators. To assess intra-rater reliability (test–retest reproducibility), 26 of the 80 Japanese participants were assessed twice at an interval of 1–3 weeks. This 1–3-week period was considered to be short enough to avoid pathophysiological progression, but long enough to reduce the chance that the subjects would recollect their previous responses.

We independently recruited a total of 41 male Japanese subjects with SBMA as the SBMAFRS-J evaluation group and followed them for 48 weeks to assess the usefulness of this outcome measure to detect disease progression. The inclusion criteria were the same as those of the SBMAFRS-J validation group.

We also recruited a total of 15 male subjects with SBMA in the USA as the SBMAFRS-E validation group. The inclusion criteria were the same as those of the Japanese version. To assess inter-rater reliability, all participants were assessed twice by two independent evaluators.

We followed the Japanese subjects at Nagoya University Hospital, and collected their data between May 2010 and July 2014. The US subjects were followed at NIH, and their data were collected between September 2012 and May 2013.

This study also included 41 Japanese male healthy controls matched for age and with no significant medical history of neurological disorders. We informed the controls that their results would be compared to those of subjects with SBMA, and they were evaluated at Kojunkai Chuo Clinic, Tokai, Japan.

This study conformed to the ethics guidelines for human genome/gene analysis research and to those for epidemiological studies endorsed by the Japanese government. The ethical committee of Nagoya University Graduate School of Medicine approved this study, and all participants gave their written informed consent. All of the US subjects signed informed consent at the National Institutes of Health in Bethesda, Maryland, USA.

### 2.3. Other clinical parameters

We used two other functional rating scales that are commonly applied to evaluate disease severity, i.e., the ALSFRS-R and the modified Norris score, to assess the external validity of the SBMAFRS.

The ALSFRS is a validated, questionnaire-based scale that measures the physical function of subjects with ALS in carrying out activities of daily living [12]. The revised version of this scale, the ALSFRS-R, was generated to improve the disproportionate weighting to limb and bulbar function, as compared to respiratory dysfunction. Each of the 12 items is rated in 4 ordinal categories, and thus the best possible score is 48. The ALSFRS-R was translated into Japanese, and its validity has been confirmed [13].

The modified Norris Score is another rating scale for ALS that consists of two parts, i.e., the Limb Norris Score and the Norris Bulbar Score. The Limb Norris Score has 21 items to evaluate limb function and the Norris Bulbar Score has 13 items to assess bulbar function. Each item is rated in 4 ordinal categories, and thus the best possible scores are 63 and 39, respectively. The original version was translated into Japanese, and its validity has been confirmed [14].

## 2.4. Statistical analysis

Statistical analyses were performed using SPSS Statistics 17.0 (SPSS Japan, Inc., Tokyo, Japan). We used descriptive variables such as the mean and standard deviation to summarize the quantitative measures. The normality of the score distribution was assessed by Shapiro–Wilk test. The score distributions were also assessed by skewness and kurtosis. For multiple comparisons, *p* values were corrected using Dunnett’s test. We examined scaling assumption (confirmatory factor analyses); acceptability (distribution of the subscales and total score); reliability (test–retest reliability: intra-class correlation coefficients [ICCs] with a one-way mixed effect model; inter-rater reliability: ICCs with a two-way mixed effect model); and validity (internal consistency: Cronbach’s alpha; inter-scale correlations; and convergent and discriminate construct validity). Often, 0.70 is recommended as a minimum standard for ICCs to indicate reliability. A value of 0.70–0.90 is considered to indicate good internal consistency for Cronbach’s alpha [15]. In the longitudinal analysis, a standardized response mean (SRM) was also calculated as an index of the effect size for a direct comparison among the outcome measures. Values of 0.20, 0.50, and 0.80 were considered to represent small, moderate, and large changes over time, respectively [16]. The SRM was calculated using the ratio of the mean score change to the standard deviation of score change.

## 3. Results

### 3.1. Items and alternatives of the SBMAFRS

The SBMAFRS consists of 14 items, each of which contains 5 (0–4) alternatives (Table 1). The 14 items consist of 5 subscales: bulbar-, upper limb-, trunk-, lower limb-, and breathing-related. The total available score ranges from 0 (worst) to 56 (normal). The bulbar-related subscale contains 5 items, the upper limb-related subscale contains 2 items, the trunk-related subscale contains 4 items, the lower limb-related subscale contains 2 items, and the breathing-related subscale contains 1 item. To minimize the difference in the questioning and examination styles used among raters, a rating algorithm that specified the exact procedures to be used was attached to this scale (Fig. SA-1).

### 3.2. Demographics of study population

The background characteristics of the subjects included in the present study were similar to those of previous clinical studies [7,8,10]. There was no significant difference in demographics among the three study populations: Scale validation group (Japan), Scale evaluation group (Japan), and Scale validation group (USA) (Table 2).

### 3.3. Score distribution

The scores of the SBMAFRS-J and SBMAFRS-E were  $36.5 \pm 7.4$  (20–54) and  $37.3 \pm 3.8$  (31–43), respectively. The total score was distributed in a normal manner (Fig. 1A). Absolute values of skewness and kurtosis of the SBMAFRS total scores were smaller than those of other functional scales, suggesting the distribution normality of SBMAFRS (Table SA-2). The comparison between the total score distribution of the SBMAFRS-J and ALSFRS-R showed that the SBMAFRS-J scores varied from low to high without a ceiling or floor effect, although the distribution of the ALSFRS-R scores was skewed to the right

(Fig. 1B). The total score of the SBMAFRS-J for the 41 healthy controls was  $55.9 \pm 0.4$  (54–56), which was close to the highest possible score.

### 3.4. Scale validation

**3.4.1. Reliability estimates**—Intra-rater agreement based on the total score of the SBMAFRS-J obtained in 26 Japanese subjects with SBMA was excellent. Detailed analysis of the ICCs for each subscale demonstrated that all subscales, except for the breathing function subscale, had an excellent intra-rater agreement (Table 3). Inter-rater reliability was also high, although detailed analysis of the ICCs for each item indicated lower values than for intra-rater reliability. Internal consistency was satisfactory, with Cronbach's alpha values ranging from 0.700 to 0.822 (Table 3). Inter-rater reliability and internal consistency of the SBMAFRS-E were similar to those of the Japanese version of this scale, SBMAFRS-J, although the ICC for the breathing-related subscale of the SBMAFRS-E was better than that of the SBMAFRS-J.

**3.4.2. Validity estimate**—As a “gold standard” disease-specific functional parameter has not been established for SBMA, we could not evaluate the criterion validity. Instead, we assessed the correlation of the subscales of the SBMAFRS with those of existing functional scales to evaluate their convergent and discriminant construct validity (Table 4). The directions and patterns of correlations were consistent with what we assumed. For example, the bulbar-related subscale of the SBMAFRS correlated best with that of the ALSFRS-R and Norris Bulbar Score, and worst with the lower limb- or respiration-related subscale of the ALSFRS-R and Limb Norris Score. Conversely, the lower limb-related subscale of the SBMAFRS correlated best with that of the ALSFRS-R and Limb Norris Score, and worst with the bulbar-related subscale of the ALSFRS-R and Norris Bulbar Score.

**3.4.3. Confirmatory factor analysis**—As stated above, the SBMAFRS was developed with the assumption that the items can be categorized into the following 5 subscales: bulbar-, upper limb-, trunk-, lower limb-, and breathing-related. To verify this assumption, we applied factor analysis using oblique varimax rotation for the SBMAFRS-J (Table SA-3). This analysis extracted 3 factors that contributed 60.0% of the total variance. Factor 1 mainly included the trunk- and lower limb-related subscales, while the bulbar- and upper limb-related subscales both contributed to factor 2. These results confirmed the factorial validity of the SBMAFRS, although the bulbar and upper limb subscales were not divided; neither were the trunk and lower limb subscales. Thus, the motor function of the SBMA subjects assessed using the SBMAFRS appears to be classified into 3 domains: bulbar and upper limb; trunk and lower limb; and respiratory. These findings were consistent with our previous study showing that bulbar impairment is closely related to upper limb dysfunction, and trunk dysfunction is associated with lower limb weakness in SBMA [10].

**3.4.4. Longitudinal analysis in the scale evaluation group**—To evaluate the sensitivity of the SBMAFRS as an outcome measure, we prospectively analyzed longitudinal changes in the SBMAFRS-J and other functional parameters for 48 weeks. The results showed slow but steady disease progression in all subscales of the SBMAFRS-J, as well as in the other functional parameters (Table 5, Fig. 2, and Fig. SA-3). To facilitate a



direct comparison of responsiveness among the various functional parameters, we calculated the SRM as an index of the effect size by the longitudinal mean change/standard deviation of score change for each parameter (Table 5). The SRM, which enables a comparison of sensitivity among the outcome measures, of the SBMAFRS was larger than that of the other scales, indicating that the SBMAFRS would require the smallest sample size in clinical trials and thus appears to be the most sensitive functional parameter [17,18]. Sample size estimation based on this longitudinal analysis was the lowest for the SBMAFRS, followed by the ALSFRS-R and Norris Bulbar Score, suggesting that the SBMAFRS was a sensitive clinical measure which detects the disease progression over time (Fig. 3).

## 4. Discussion

Outcome measures that enable the rigorous quantification of disease severity or quality of life are essential to estimate the efficacy of potential therapies [19–24]. Disease-specific functional rating scales that measure global dysfunction, based on a structured psychometric validation procedure, are increasingly important in defining the primary endpoint in clinical trials, because the quality of the rating scales has the potential to determine the validity of such studies [25–27]. Nevertheless, to date, no disease-specific outcome measures for SBMA are available. In the present study, we created a disease-specific functional rating scale for SBMA (SBMAFRS) and validated this score in cross-sectional and longitudinal analysis. One major advantage of the SBMAFRS is that it was developed as a compound scale to capture the multiple aspects of SBMA. The results of the present study suggest that the SBMAFRS was valid and reliable measure for the quantitative evaluation of SBMA. In the longitudinal evaluation, the change of the SBMAFRS-J over time was more readily detectable than that of the other functional parameters, although still larger sample sizes are required for detection of therapeutic efficacy. These results demonstrate that the SBMAFRS is a disease-specific outcome measure that can be used for clinical studies.

The SBMAFRS was designed to overcome several problems that arose when we applied the functional measures for ALS or other neuromuscular disorders to SBMA. First, it tends to be difficult to notice the chronological changes of neurological symptoms in subjects with SBMA because disease progression in this disease is much slower than in ALS. In the process of scale development, each item was designed to detect subtle symptoms by subdividing the alternatives of the milder symptoms. Second, the vague descriptions of the alternatives in the existing clinical measures might result in different responses between raters. We developed each alternative of the SBMAFRS to ensure that the expressions used were as clear as possible. As a result, the scores of the SBMAFRS were distributed in a wider range than those of the ALSFRS-R, suggesting that the SBMAFRS measures the physical function of individual subjects with mild motor dysfunction more sensitively than the ALSFRS-R.

The confirmatory factor analysis of this scale indicated that the bulbar function is strongly associated with the upper limb function in spinal and bulbar muscular atrophy, as we previously demonstrated [10]. A similar relationship is also shown in ALS: the degree of bulbar palsy is more strongly associated with upper limb weakness than lower limb weakness [28]. As a possible pathomechanism for this phenomenon, spread of lesions from

the bulbar region to the upper limb is demonstrated in ALS [29]. Given that the spread of neuropathology is also suggested in polyglutamine diseases, a similar mechanism may underlie the pathophysiology of SBMA [30].

The SBMAFRS-J was also examined prospectively over 48 weeks to evaluate its sensitivity to changes over time. To this end, we calculated the SRM as the index of the effect size. Calculation of the SRM enables a direct comparison of the sensitivity of the scales to detect changes among outcome measures. Compared to the other measures, the SBMAFRS had a larger SRM, indicating that its use might reduce the number of subjects needed for clinical trials. The effect sizes of disease-modifying therapies in slowly progressive neurodegenerative disorders generally tend to be small because of smaller longitudinal changes in proportion to large symptomatic daily fluctuations [31–33]. Although a relatively small SRM is an inevitable feature of slowly progressive neurodegenerative diseases such as SBMA, effort should be made to further increase the sensitivity of this scale through the experience gained by its use in future clinical trials.

Although our results suggest that the SBMAFRS is a valid disease-specific score of SBMA, there are limitations that need to be improved. First, low reliability of the breathing-related item is a critical drawback of the SBMAFRS-J, which is attributable to several factors. One of the most important issues was that each alternative of breathing-related item was categorized within a narrow range, because respiratory function in the most of patients with SBMA is preserved and a few need ventilation support [5]. This over-categorization might lead to the discrepancies among evaluations and result in low intra-rater or inter-rater agreement. In addition, a minor disagreement led to low ICC for the breathing-related item of the SBMAFRS-J, since the score distribution biased extremely: 58 out of 80 subjects were judged as having the same score in the inter-rater agreement assessment (Table SA-4A). Similar results were also obtained as for the intra-rater reproducibility (Table SA-4B). Furthermore, the number of raters was eight for the SBMAFRS-J, but two for SBMAFRS-E. This difference might also lead to the differential reliability of breathing-related subscale of the SBMAFRS-E, since only one item was included in this subscore. The small number of the US subjects and relatively short longitudinal follow-up period also limit the interpretation of our results. In the future study, a larger cohort needs to be followed-up for longer time to confirm the utility of the SBMAFRS. Second, the old-fashioned clinimetrics that adopted items without standardization or weighing as in SBMAFRS would be disadvantageous for being used in clinical trials because of the limited linearity. In this regard, the Rasch and item response theory is an important approach to transform ordinal scores into interval measures that are scale-independent and accurate for patients' assessment [34]. In fact, this methodology has been applied for several functional scales of neurological disorders such as the 9-item fatigue severity scale (FSS) for immune-mediated neuropathies [35] and the modified Medical Research Council grading system [36]. Furthermore, the concept of minimal clinically important difference (MCID) should also be considered as a means of assessing treatment response in clinical trials [37].

In summary, we created a disease-specific functional rating scale for SBMA, a slowly progressive neurodegenerative disorder. Although our results indicate the validity and sensitiveness of this scale, the SBMAFRS, the feasibility of this scale as an outcome



measure needs to be re-assessed in interventional studies with a larger sample size including the English-speaking subjects.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### Funding

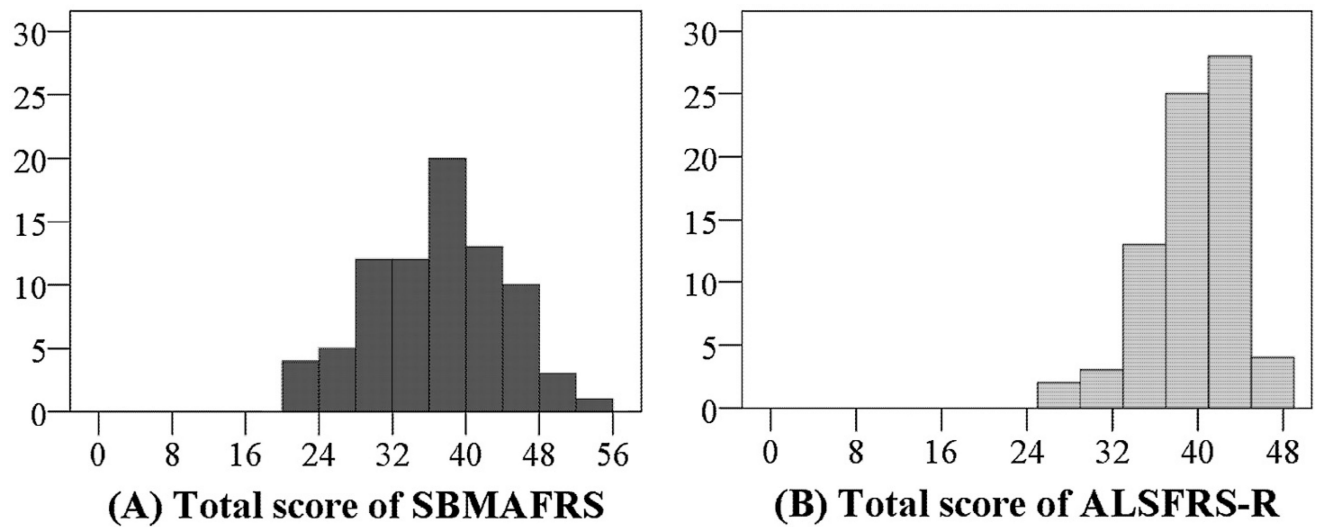
Dr. Katsuno is supported by a Grant-in-Aid for Scientific Research on Innovated Areas “Foundation of Synapse and Neurocircuit Pathology” from MEXT, Japan, (22110005); KAKENHI grants from MEXT/JSPS, Japan, (No. 22110005, 26293206, 26670440, and 26670439); and Core Research for Evolutional Science and Technology (CREST) from the Japan Science and Technology Agency (JST); and a grant from the Daiichi Sankyo Foundation of Life Science.

Dr. Sobue is supported by KAKENHI grants from MEXT/JSPS, Japan (No. 21229011); grants from the Ministry of Health, Labour and Welfare, Japan; and Core Research for Evolutional Science and Technology (CREST) from the Japan Science and Technology Agency (JST).

## References

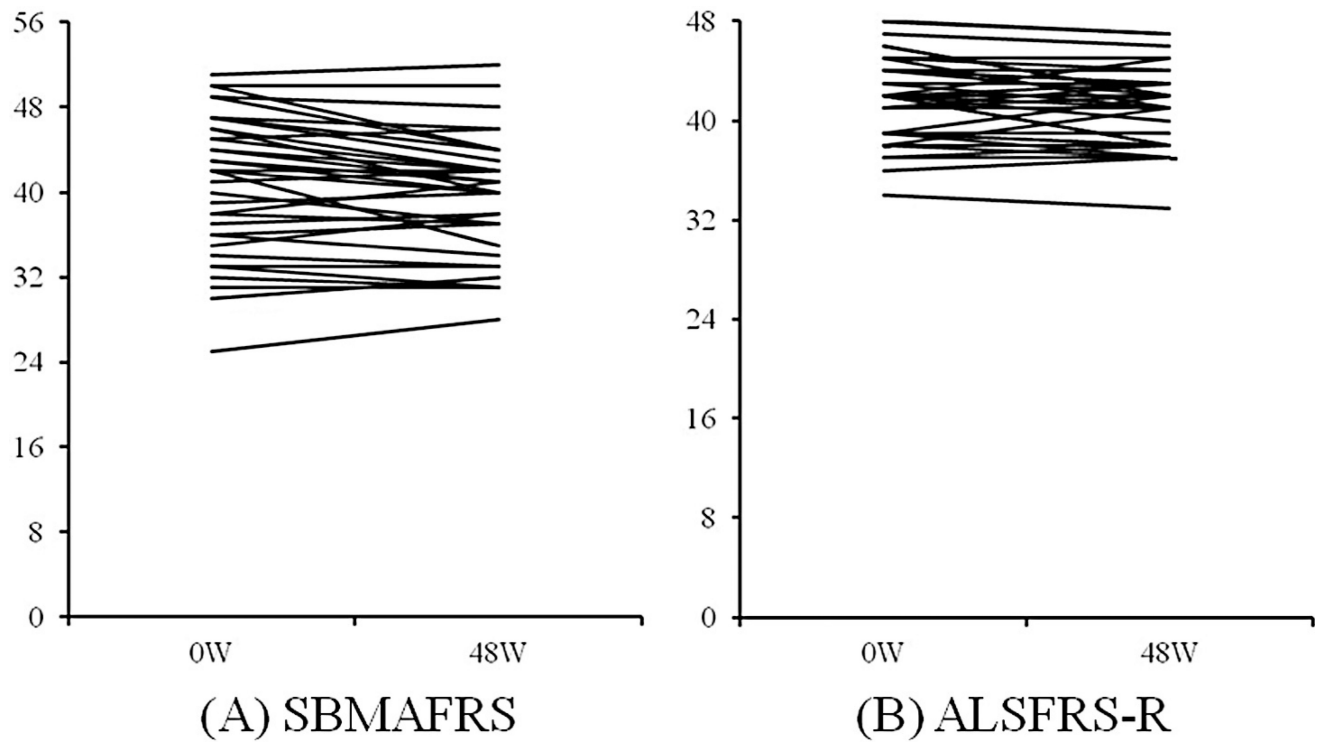
1. Kennedy WR, Alter M, Sung JH. Progressive proximal spinal and bulbar muscular atrophy of late onset. A sex-linked recessive trait. *Neurology*. 1968; 18:671–80. [PubMed: 4233749]
2. Sobue G, Hashizume Y, Mukai E, et al. X-linked recessive bulbospinal neuronopathy. A clinicopathological study. *Brain*. 1989; 112:209–32. [PubMed: 2917278]
3. Katsuno M, Tanaka F, Adachi H, et al. Pathogenesis and therapy of spinal and bulbar muscular atrophy (SBMA). *Prog Neurobiol*. 2012; 99(3):246–56. [PubMed: 22609045]
4. Sperfeld AD, Karitzky J, Brummer D, et al. X-linked bulbospinal neuronopathy: Kennedy disease. *Arch Neurol*. 2002; 59:1921–6. [PubMed: 12470181]
5. Atsuta N, Watanabe H, Ito M, et al. Natural history of spinal and bulbar muscular atrophy (SBMA): a study of 223 Japanese patients. *Brain*. 2006; 129:1446–55. [PubMed: 16621916]
6. Banno H, Katsuno M, Suzuki K, et al. Phase 2 trial of leuprorelin in patients with spinal and bulbar muscular atrophy. *Ann Neurol*. 2009; 65:140–50. [PubMed: 19259967]
7. Katsuno M, Banno H, Suzuki K, et al. Efficacy and safety of leuprorelin in patients with spinal and bulbar muscular atrophy (JASMITT study): a multicentre, randomised, double-blind, placebo-controlled trial. *Lancet Neurol*. 2010; 9:875–84. [PubMed: 20691641]
8. Fernández-Rhodes LE, Kokkinis AD, White MJ, et al. Efficacy and safety of dutasteride in patients with spinal and bulbar muscular atrophy: a randomised placebo-controlled trial. *Lancet Neurol*. 2011; 10:140–7. [PubMed: 21216197]
9. Querin G, D’Ascenzo C, Peterle E, et al. Pilot trial of clenbuterol in spinal and bulbar muscular atrophy. *Neurology*. 2013; 80:2095–8. [PubMed: 23645595]
10. Hashizume A, Katsuno M, Banno H, et al. Longitudinal changes of outcome measures in spinal and bulbar muscular atrophy. *Brain*. 2012; 135:2838–48. [PubMed: 22773541]
11. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993; 46:1417–32.
12. The ALS CNTF Treatment Study (ACTS) Phase I-II Study Group. The amyotrophic lateral sclerosis functional rating scale. Assessment of activities of daily living in patients with amyotrophic lateral sclerosis. *Arch Neurol*. 1996; 53:141–7. [PubMed: 8639063]
13. Ohashi Y, Tashiro K, Itoyama Y, et al. Study of functional rating scale for amyotrophic lateral sclerosis: revised ALSFRS(ALSFRS-R) Japanese version. *No to Shinkei*. 2001; 53:346–55. [PubMed: 11360474]
14. Oda E, Ohashi Y, Tashiro K, et al. Reliability and factorial structure of a rating scale for amyotrophic lateral sclerosis. *No to Shinkei*. 1996; 48:999–1007. [PubMed: 8951891]

15. Nunnally, JC., Bemstein, IH. Psychometric theory. 3. New York: McGraw-Hill; 1994.
16. Kazis LE, Anderson JJ, Meenan RF. Effect sizes for interpreting changes in health status. *Med Care*. 1989; 27(3 Suppl):S178–89. [PubMed: 2646488]
17. Giordano A, Pucci E, Naldi P, et al. Responsiveness of patient reported outcome measures in multiple sclerosis relapses: the REMS study. *J Neurol Neurosurg Psychiatry*. 2009; 80:1023–8. [PubMed: 19443471]
18. Schmitz-Hübsch T, Fimmers R, Rakowicz M, et al. Responsiveness of different rating instruments in spinocerebellar ataxia patients. *Neurology*. 2010; 74:678–84. [PubMed: 20177122]
19. Burns J, Ouvrier R, Estilow T, et al. Validation of the Charcot-Marie-Tooth disease pediatric scale as an outcome measure of disability. *Ann Neurol*. 2012; 71:642–52. [PubMed: 22522479]
20. Simmons Z, Felgoise SH, Bremer BA, et al. The ALSSQOL. Balancing physical and nonphysical factors in assessing quality of life in ALS. *Neurology*. 2006; 67:1659–64. [PubMed: 17101900]
21. Shy ME, Blake J, Krajewski K, et al. Reliability and validation of the CMT neuropathy score as a measure of disability. *Neurology*. 2005; 64:1209–14. [PubMed: 15824348]
22. Müller J, Wissel J, Kemmler G, et al. Craniocervical dystonia questionnaire (CDQ-24): development and validation of a disease-specific quality of life instrument. *J Neurol Neurosurg Psychiatry*. 2004; 75:749–53. [PubMed: 15090572]
23. Payan CA, Viallet F, Landwehrmeyer BG, et al. Disease severity and progression in progressive supranuclear palsy and multiple system atrophy: validation of the NNIPPS-Parkinson Plus Scale. *PLoS ONE*. 2011; 6(8):e22293. [PubMed: 21829612]
24. Hashizume A, Katsuno M, Banno H, et al. Difference in chronological changes of outcome measures between untreated and placebo-treated patients of spinal and bulbar muscular atrophy. *J Neurol*. 2012; 259:712–19. [PubMed: 21952871]
25. Huntington Study Group. Unified Huntington's disease rating scale: reliability and consistency. *Mov Disord*. 1996; 11:136–42. [PubMed: 8684382]
26. Wenning GK, Tison F, Seppi K, et al. Development and validation of the Unified Multiple System Atrophy Rating Scale (UMSARS). *Mov Disord*. 2004; 19:1391–402. [PubMed: 15452868]
27. Hobart JC, Cano SJ, Zajicek JP, et al. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol*. 2007; 6:1094–105. [PubMed: 18031706]
28. Simon NG, Lomen-Hoerth C, Kiernan MC. Patterns of clinical and electrodiagnostic abnormality in early amyotrophic sclerosis. *Muscle Nerve*. 2014; 50:894–9. [PubMed: 24647900]
29. Ravits JM, La Spada AR. ALS motor phenotype heterogeneity, focality, and spread: deconstructing motor neuron degeneration. *Neurology*. 2009; 73:805–11. [PubMed: 19738176]
30. Cicchetti F, Lacroix S, Cisbani G, et al. Mutant huntingtin is present in neuronal grafts in Huntington disease patients. *Ann Neurol*. 2014; 76:31–42. [PubMed: 24798518]
31. Marder K, Zhao H, Myers RH, et al. Rate of functional decline in Huntington's disease. Huntington Study Group. *Neurology*. 2000; 54:452–8. [PubMed: 10668713]
32. Martinez-Martin P, Prieto L, Forjaz MJ. Longitudinal metric properties of disability rating scales for Parkinson's disease. *Value Health*. 2006; 9:386–93. [PubMed: 17076869]
33. Martínez-Martin P, Carod-Artal FJ, da Silveira Ribeiro L, et al. Longitudinal psychometric attributes, responsiveness, and importance of change: an approach using the SCOPA-Psychosocial questionnaire. *Mov Disord*. 2008; 23:1516–23. [PubMed: 18649392]
34. Merkies IS, Faber CG. Outcome measures in Duchenne muscular dystrophy: are we ready for the new therapeutic era? *Neuromuscul Disord*. 2009; 19:447. [PubMed: 19482474]
35. Draak TH, Vanhoutte EK, van Nes SI, et al. Changing outcome in inflammatory neuropathies: Rasch-comparative responsiveness. *Neurology*. 2014; 83:2124–32. [PubMed: 25378677]
36. Vanhoutte EK, Faber CG, van Nes SI, et al. Modifying the Medical Research Council grading system through Rasch analyses. *Brain*. 2012; 135:1639–49. [PubMed: 22189568]
37. Smith AG, Burns TM. Reevaluating clinical measurement tools in therapeutic trials: time to make a Rasch decision? *Neurology*. 2014; 83:2104–5. [PubMed: 25378676]

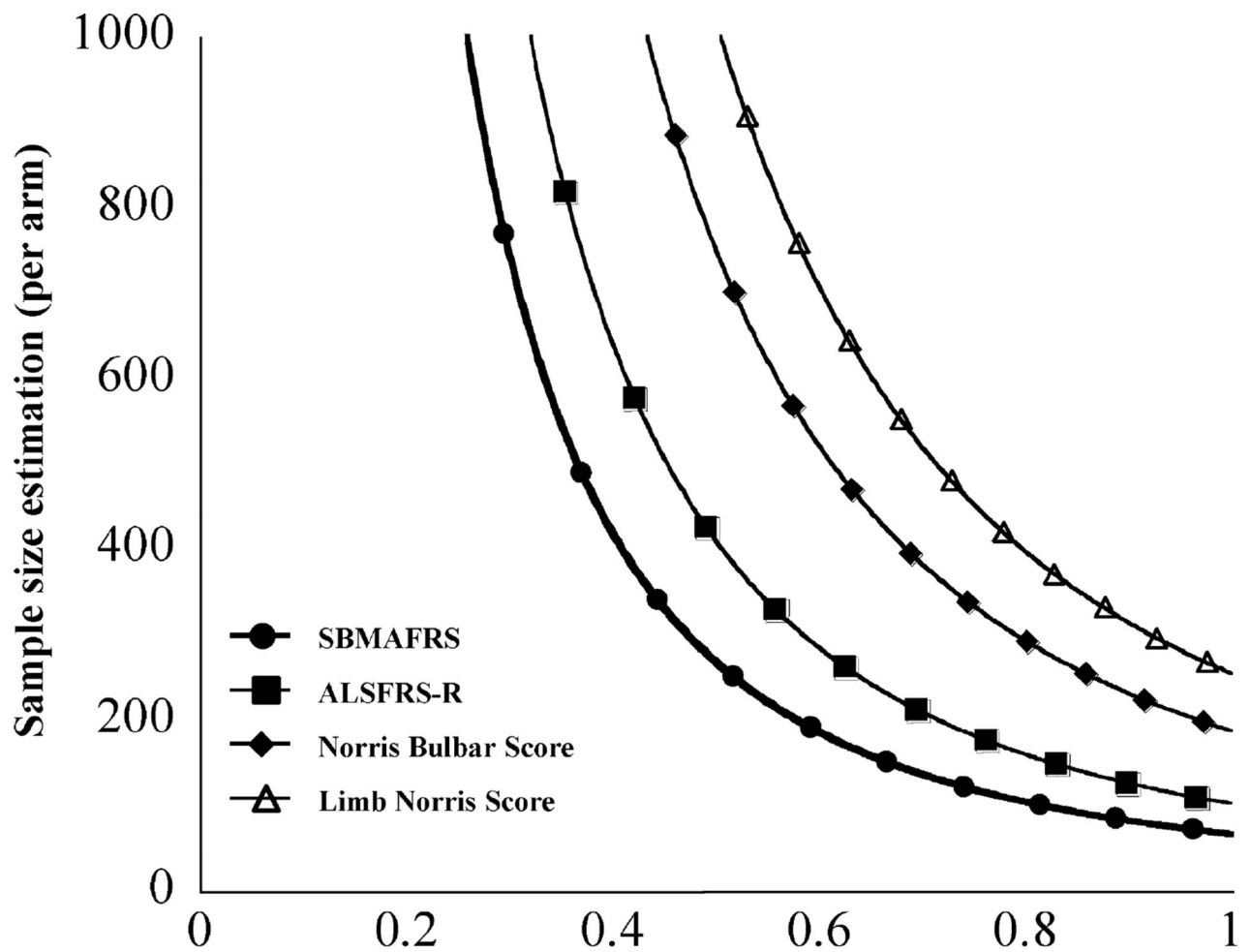


**Fig. 1.**

Score distribution of the SBMAFRS and the ALSFRS-R. The SBMAFRS-J was distributed in a normal manner (A). The total score distribution of the SBMAFRS varied from low to high without ceiling or floor effects (A), although that of the ALSFRS-R skewed to the right with a ceiling effect (B).



**Fig. 2.**  
Individual change of the functional scores. The change of the total score of SBMAFRS (A)  
and that of ALSFRS-R (B) in individual subjects.



### Expected reduction of the natural progression rate with intervention

**Fig. 3.**

Sample size estimation. Estimated sample size per arm for a 2-arm interventional clinical trial based on the observed variability of the chronological progression rate found in the present study (power 0.80,  $\alpha$ 0.05). SBMAFRS = Spinal and Bulbar Muscular Atrophy Functional Rating Scale; ALSFRS-R = Revised Amyotrophic Lateral Sclerosis Functional Rating Scale.

**Table 1****Spinal and bulbar muscular atrophy functional rating scale (SBMAFRS).**

<b>Bulbar-related subscale</b>	
Speech	4. Normal speech. 3. Speaks in a slightly nasal tone. 2. What is said can be understandable despite impaired speech. 1. A portion of what is said cannot be understandable unless heard over and over again. 0. Other media than voices are needed for communication.
Control of salivation	4. No drooling. 3. Dribbles at night or during eating (less often than once a week). 2. Occasionally dribbles at night or during eating (not less often than once a week). 1. Dribbles on other occasions than above. 0. Dribbles so much and thus always needs tissues or a handkerchief.
Swallowing	4. No difficulty. 3. Occasionally chokes on food depending on its texture. 2. Avoids some particular textures of food that may cause choking, etc. 1. Food textures have to be always modified to soft or chopped ones. 0. Enteral or intravenous nutrition is needed.
Tongue	4. No atrophy. 3. Atrophic, but the tongue tip can be smoothly moved to and fro between the left and right corners of the mouth. 2. The tongue tip can be moved, but not smoothly, to and fro between the left and right corners of the mouth. 1. The tongue tip cannot reach the mouth corner. 0. No motion of the tongue.
Puffing cheeks	4. The air does not leak when the cheeks are firmly pressed. 3. The air leaks when the cheeks are firmly pressed. 2. The cheeks can be puffed out, but not fully. 1. Although the lips can be pursed, the cheeks cannot be puffed. 0. The lips cannot be pursed.
<b>Upper limb-related subscale</b>	
Writing	4. No difficulty. 3. Slightly difficult in writing, but can write a long sentence as well. 2. Writing gets slow in writing a long sentence. 1. Difficult in writing even a short sentence. 0. Unable to write letters.
Eating action	4. No problem with eating actions. 3. Able to finish eating within a normal time, but has some difficulty in using a knife and a fork. 2. Takes a longer eating time due to his difficulty in using a knife and a fork. 1. Unable to use a knife or a fork, but good enough at using a spoon. 0. Needs assistance for eating.
<b>Trunk-related subscale</b>	
Dressing activity	4. No problem. 3. Partially impaired, but taking no additional time to change clothes.



	2. Partially impaired, so takes time to change clothes because he needs to sit or lean against something.
	1. Partial assistance is needed.
	0. Total assistance is needed.
Rising from a sitting position	4. No problem.
	3. Minor problem.
	2. Able to rise all alone, but needs to hold on something.
	1. Any assistance is needed.
	0. Unable to rise at all.
Arising from a supine position	4. No problem.
	3. Minor problem.
	2. Able to arise all alone, but needs some efforts, ex., holding a lateral position beforehand.
	1. Any assistance is needed to arise.
	0. Unable to arise at all.
Bowing	4. Able to bow his head fully in a standing position.
	3. Able to bow his head in a standing position, but not fully.
	2. Able to hold a standing position, but slightly hard to bow his head.
	1. Difficult in holding a standing position, but able to bow his head in a sitting position.
	0. Unable to bow his head even in a sitting position.
Lower limb-related subscale	
Walking	4. No problem with walking or running.
	3. No problem with walking, but it is difficult to run.
	2. A cane or some other assistive device is sometimes needed.
	1. A cane or some other assistive device is always needed.
	0. Walking is impossible.
Stairs	4. No problem with walking upstairs.
	3. Able to walk upstairs without a handrail, but slowly.
	2. Able to walk upstairs without assistance if a handrail is available.
	1. Assistance from someone is needed.
	0. Unable to walk upstairs anyhow.
Breathing-related subscale	
Breathing	4. No dyspnea
	3. Gets short of breath during activity like walking.
	2. Gets short of breath even during slight activity like standing up.
	1. Has shortness of breath even at rest.
	0. A respiratory assistor such as NIPPV and TPPV is needed.

**Table 2**

Clinical profiles of the Japanese and US patients with SBMA.

	Scale validation group (Japan) (n = 80)	Scale evaluation group (Japan) (n = 41)	Scale validation group (USA) n = 15)	P
Age at first evaluation (years)	57.7 ± 10.0 (37–82)	51.9 ± 11.2 (30–73)	56.1 ± 10.9 (32–46)	NS
Disease duration (years)	14.7 ± 7.3 (1–37)	9.8 ± 5.7 (1–24)	17.5 ± 12.2 (3–49)	NS
Age at onset (years)	43.1 ± 11.4 (18–66)	41.7 ± 11.8 (22–66)	38.6 ± 13.5 (18–60)	NS
CAG repeat length on AR gene	48.3 ± 3.7 (40–60)	49.6 ± 4.8 (40–60)	46.7 ± 2.6 (43–51)	NS
SBMAFRS total score	36.5 ± 7.4 (20–54)	41.3 ± 6.9 (25–51)	37.3 ± 3.8 (31–43)	NS
ALSFRS-R total score	39.1 ± 4.2 (27–47)	41.8 ± 3.4 (34–48)	37.7 ± 3.2 (32–46)	NS
Creatine kinase (IU/L)	781 ± 653 (103–4164)	1189 ± 763 (249–2296)	1119 ± 512 (491–2296)	NS
Creatinine (mg/dL)	0.41 ± 0.13 (0.19–0.79)	0.46 ± 0.08 (0.30–0.72)	0.51 ± 0.11 (0.33–0.72)	NS
Testosterone (µg/dL)	5.12 ± 3.82 (0.20–18.42)	7.01 ± 2.80 (2.77–13.82)	4.38 ± 1.58 (1.47–7.34)	NS

SBMAFRS = spinal and bulbar muscular atrophy functional rating scale; ALSFRS-R = revised amyotrophic lateral sclerosis functional rating scale; NS = not significant.

Each cell contains the mean ± SD.

**Table 3**

Intra- and inter-rater reliability and internal consistency of each subscale in the SBMAFRS.

		<b>Bulbar</b>	<b>Upper limbs</b>	<b>Trunk</b>	<b>Lower limbs</b>	<b>Breathing</b>	<b>Total score</b>
<b>Reliability</b>							
Intra-rater agreement (95% CI)	(Japanese)	0.890	0.828	0.935	0.957	0.076	0.910
	(n = 26)	(0.772–0.949)	(0.656–0.919)	(0.863–0.970)	(0.907–0.980)	(–0.309–0.442)	(0.811–0.958)
Inter-rater agreement (95% CI)	(Japanese)	0.658	0.745	0.817	0.831	0.297	0.797
	(n = 80)	(0.477–0.781)	(0.602–0.836)	(0.715–0.883)	(0.737–0.892)	(0–0.549)	(0.684–0.870)
	(US)	0.808	0.665	0.786	0.823	0.985	0.837
Internal consistency	(n = 15)	(0.428–0.936)	(0.004–0.888)	(0.361–0.928)	(0.473–0.941)	(0.955–0.995)	(0.515–0.945)
<b>Internal consistency</b>							
Cronbach's alpha	(Japanese)	0.700	0.712	0.822	0.784	NA	NA
	(n = 80)						
	(US)	0.438	0.393	0.519	0.494	NA	NA
	(n = 15)						

NA = not applicable; CI = confidence interval.

Inter-rater and intra-rater agreement was evaluated by using the intra-class correlation coefficient.

**Table 4**

Convergent and discriminant construct validity of the SBMAFRS-J.

SBMAFRS-J						
	Bulbar	Upper limbs	Trunk	Lower limbs	Breathing	Total score
ALSFRS-R	0.873	0.542	0.257	0.196	0.222	0.635
	$p < 0.001$	$p < 0.001$	$p = 0.042$	$p = 0.124$	$p = 0.080$	$p < 0.001$
Upper limbs	0.549	0.856	0.469	0.432	0.355	0.714
	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.004$	$p < 0.001$
Trunk	0.354	0.436	0.741	0.747	0.279	0.734
	$p = 0.004$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.027$	$p < 0.001$
Lower limbs	0.186	0.385	0.834	0.852	0.312	0.693
	$p = 0.145$	$p = 0.002$	$p < 0.001$	$p < 0.001$	$p = 0.013$	$p < 0.001$
Respiration	0.043	0.115	0.226	0.268	0.662	0.247
	$p = 0.736$	$p = 0.367$	$p = 0.075$	$p = 0.034$	$p < 0.001$	$p = 0.051$
Limb Norris Score	0.370	0.578	0.847	0.849	0.366	0.853
	$p = 0.003$	$p < 0.001$	$p < 0.001$	$p < 0.001$	$p = 0.003$	$p < 0.001$
Norris Bulbar Score	0.801	0.511	0.365	0.292	0.181	0.680
	$p < 0.001$	$p < 0.001$	$p = 0.003$	$p = 0.020$	$p = 0.156$	$p < 0.001$
Age at examination	0.019	0.032	-0.269	-0.179	-0.161	-0.145
	$p = 0.879$	$p = 0.797$	$p = 0.027$	$p = 0.144$	$p = 0.189$	$p = 0.238$
Disease duration	0.100	-0.140	-0.347	-0.427	-0.095	-0.240
	$p = 0.419$	$p = 0.255$	$p = 0.004$	$p < 0.001$	$p = 0.442$	$p = 0.049$

SBMAFRS = spinal and bulbar muscular atrophy functional rating scale; ALSFRS-R = revised amyotrophic lateral sclerosis functional rating scale.

Each cell contains Spearman's correlation coefficient (upper line) and the corresponding  $p$  value (lower line).

**Table 5**

Longitudinal analyses of the SBMAFRS-J and other functional parameters.

N = 41	Baseline (SD)	Follow-up (SD)	Mean longitudinal change (SE)	SRM
SBMAFRS				
Bulbar	14.5 (3.1)	14.4 (2.9)	−0.22 (0.25)	−0.14
Upper limbs	6.0 (1.7)	5.7 (1.6)	−0.27 (0.15)	−0.29
Trunk	12.0 (2.5)	11.3 (2.6)	−0.66 (0.22)	−0.49
Lower limbs	5.0 (1.4)	4.9 (1.2)	−0.10 (0.15)	−0.11
Breathing	3.8 (0.4)	3.7 (0.5)	−0.02 (0.09)	−0.05
Total score	41.3 (6.9)	40.0 (6.2)	−1.27 (0.44)	−0.48
ALSFRS-R	41.8 (3.4)	41.0 (3.2)	−0.73 (0.31)	−0.39
Limb Norris Score	55.7 (5.3)	55.1 (5.3)	−0.56 (0.38)	−0.25
Norris Bulbar Score	33.3 (4.8)	32.6 (4.7)	−0.73 (0.42)	−0.29

SBMAFRS = spinal and bulbar muscular atrophy functional rating scale; ALSFRS-R = revised amyotrophic lateral sclerosis functional rating scale; SD = standard deviation; SE = standard error of the mean; SRM = standardized response mean.