Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Development and longitudinal evaluation of ALSFRS-R-SE version 2.1 for remote digital self-assessment

André Maier

andre.maier@charite.de

Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

Matthias Boentert

Neuromedical Centre, Department of Sleep Medicine, Klinikum Osnabrück

Peter Reilich

Department of Neurology with Friedrich-Baur-Institute, University Hospital, Ludwig-Maximilians-University (LMU)

Torsten Grehl

Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Alfried Krupp Krankenhaus

Janina von der Gablentz

University of Luebeck, Department of Neurology, Precision Neurology of Neuromuscular and Motor Neuron Diseases

Paul Lingor

Department of Neurology, School of Medicine and Health, TUM University Hospital Rechts der Isar, Technical University of Munich

Bertram Walter

Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

Susanne Spittel

Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

Peggy Schumann

Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu

Berlin

Senthil Subramanian
Ambulanzpartner Soziotechnologie APST GmbH

Alessio Riitano
Ambulanzpartner Soziotechnologie APST GmbH

Christoph Münch
Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

Thomas Meyer
Department of Neurology, Center for ALS and Other Motor Neuron Disorders, Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin and Humboldt-Universität zu Berlin

---

Research Article

Keywords: Amyotrophic lateral sclerosis, ALSFRS-R-SE, self-assessment, patient-reported outcome, remote digital assessment, ALSFRS-R

Posted Date: November 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11420195/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: Competing interest reported. T.M. and C.M. are founders and shareholders of the Ambulanzpartner Soziotechnologie APST GmbH, which makes the mobile application “ALS App”. All other authors have no competing interests to declare that are relevant to the content of this article.

Abstract

Background

The self-explanatory version of the Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-R-SE) was developed to enable standardized assessment without a trained interviewer (version 1.0). However, it needed further refinement to better support longitudinal remote digital self-assessment.

Methods

Following a structured, consensus-based revision in both German and English, we analyzed remote assessments from the ALS App collected between March 2024 and July 2026. Using linear regression, we estimated the slope of the total score before the transition, the immediate change, and the change in slope for 189 participants. In the predictive analyses, the effects on 12 items and 48 response boundaries were evaluated for 420 participants.

Results

The combined word count decreased by 16.0% in German and by 23.6% in English. The reductions were particularly pronounced in the introductory text and in the response descriptions. V2.1 also showed structural simplification, including fewer parenthetical insertions. At the time of the transition, the total score change was − 0.062 points (95% CI − 0.303 to 0.184) and the change in slope was + 0.186 points/month (− 0.054 to 0.420). The item effects ranged from − 0.040 to + 0.036 points, with all confidence intervals containing the value zero. None of the 48 response boundary effects remained significant after correction for the false discovery rate.

Conclusion

ALSFRS-R-SE version 2.1 provides more concise wording and substantially reduces textual burden. Longitudinal analyses supported measurement continuity, showing no detectable systematic discontinuity in total or item scores and no response-boundary effects across the observed V1.0 to V2.1 transition in remote digital self-assessment.

INTRODUCTION

The Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised (ALSFRS-R) is the most widely used instrument for assessing functional impairment and disease progression in amyotrophic lateral sclerosis (ALS) [1]. Originally developed for clinical research, it is now used in routine care and enables monitoring of functional changes in bulbar, fine motor, gross motor, and respiratory functions. The ALSFRS-R was originally developed for administration by a trained evaluator. Although its reliability is high under

standardized conditions [2], the assessment remains influenced by how individuals interpret response options and by the consistency of scoring procedures. To reduce these variations, harmonized standard operating procedures and assessment training were introduced [3]. Nevertheless, longitudinal analyses have identified sudden increases in ALSFRS-R total scores, with their frequency varying across trial centers [4]. Evaluator performance on ALSFRS-R scoring exercises has also been associated with the mode and interval of training [5].

Self-assessment offers a complementary method that reduces reliance on different raters and allows assessments to be conducted outside of scheduled clinical visits [6, 7]. Consequently, the approach of remote digital assessment of the ALSFRS-R has been extensively studied and evaluated for its validity [8–10]. Furthermore, state-of-the-art smartphone-based assessments enable more frequent reports of functional changes, which could increase data density between visits [9, 11]. This is particularly important in ALS, where progressive motor impairment can make regular visits to specialized centers increasingly difficult for patients. In addition, remote digital assessment tools are easy to use even on small screens and can be completed with minimal physical and cognitive effort. Thus, more frequent collection of patient-reported data can be achieved without disproportionately increasing the burden on respondents [12].

In 2022, the self-explanatory version of the ALSFRS-R was developed through a structured consensus process to enable a standardized assessment without the need for a trained interviewer [13]. The ALSFRS-R-SE includes explanations and detailed descriptions of the individual response categories, while retaining the original structure of the ALSFRS-R. The scale was designed for patients, caregivers, and healthcare professionals, and for the use in both paper-and-pencil and digital formats. Subsequent studies confirmed the feasibility of remote digital assessments using the ALSFRS-R-SE and showed that the longitudinal variability was not inferior to data collected in a clinical setting [9, 14]. These findings confirmed that the ALSFRS-R-SE is suitable for independent remote functional self-assessment. As practical use has increased, particularly in digital settings, certain needs for further refinement have become apparent, including concise wording, clear presentation, and a more consistent structure of annotations across the items and response options [15, 16].

Particularly when adapting the instrument across different languages, a consistent structure and comparable content should be maintained while preserving conceptual equivalence rather than relying on literal translation [17]. Recent international initiatives have called for global standardization in implementing the ALSFRS-R emphasizing the need for a uniform format in content, structure, and language even more apparent [18, 19].

Based on these considerations, a revised version of the ALSFRS-R-SE was developed to provide a more concise, consistent, and linguistically aligned self-assessment while preserving the scale's structure and content. The present study evaluates its implementation using longitudinal data to assess whether functional assessments remain consistent during the transition from the previous version 1.0 to version 2.1.

METHODS

Study design and objectives

We first conducted a structured, consensus-based, ALS expert-led revision of the ALSFRS-R-SE from Version 1.0 (V1.0) to Version 2.1 (V2.1), followed by an observational longitudinal analysis of pseudonymized digital remote assessments collected before and after the version transition. The ALSFRS-R-SE consists of 12 items scored on a scale of 0 to 4, resulting in a total score ranging from 0 to 48, with higher scores indicating better functional status. Its original development and its use for digital remote self-assessment have been described previously [13, 14]. The exploratory longitudinal analysis is reported in accordance with the STROBE reporting guidelines [20].

Development of ALSFRS-R-SE V2.1

The aim of the revision was to reduce text length and improve linguistic clarity and suitability for digital use, while retaining the clinical content, item structure, and scoring logic of ALSFRS-R-SE Version 1.0. An initial linguistic and editorial revision produced a working draft designated V2.0, which served as the basis for the subsequent expert review. All proposed changes were reviewed by a working group, which comprised 8 ALS clinicians and researchers with expertise in ALS and the ALSFRS-R-SE from 6 ALS centers. Between November 2025 and April 2026, each of the 12 items and their corresponding response categories and annotations were reviewed systematically. The wording was iteratively refined, with particular attention to clarity, brevity, preservation of clinical meaning and consistency with the original ALSFRS-R response categories. Proposed modifications were discussed until a consensus was reached within the working group. The resulting consensus version was designated V2.1. The German and English versions of ALSFRS-R-SE V2.1 were developed and aligned in parallel throughout the entire revision process.

Data source and Ethical Approval

The longitudinal analysis was based on pseudonymized ALSFRS-R-SE assessments recorded through the ALS App (https://www.ambulanzpartner.de/als-app/) [9, 14] between March 2024 and July 2026 (Ambulanzpartner APST GmbH, Berlin, Germany). A valid assessment required complete scores ranging from 0 to 4 for all 12 items. After data cleaning, 17,025 valid data sets remained. All patients provided written consent for the use of their data for scientific purposes. Data collection was approved by the Medical Ethics Committee of Charité – Universitätsmedizin Berlin, Germany, as both a patient registry (2026-101-CCM) and an observational trial (EA1/219/15).

Patient cohorts

Patients were eligible if they were enrolled in the aforementioned registry or trial, both of which required a diagnosis of ALS according to the Gold Coast criteria [21]. Two related cohorts were defined according to the requirements of the respective analyses. The total-score analysis included 189 participants with four V1.0 assessments preceding the first V2.1 assessment and three consecutive V2.1 assessments from the transition onward, resulting in 1,323 assessments.

The item-level analysis used a broader cohort. Participants were eligible if they had at least three V1.0 assessments spanning at least 60 days, followed by a first V2.1 assessment, with all V1.0 assessments preceding all V2.1 assessments. This criterion was met by 450 participants. Of these, 420 had sufficient longitudinal V1.0 data to test the prediction method on at least one V1.0 assessment before applying it to the first V2.1 assessment. This test was used to estimate each participant's prediction error.

Statistical Methods

Sociodemographic and clinical characteristics were presented as counts and percentages for categorical variables and as means with standard deviations or medians with interquartile ranges for quantitative variables. For the textual comparison of the ALSFRS-R-SE versions, the words in each component (introductory text, response labels, response descriptions, and additional item text) were counted and compared as percentages. Formula-based readability was assessed using the Flesch Reading Ease (FRE) score, which accounts for average sentence length and average word length in syllables, with higher scores indicating a lower estimated reading effort [22]. The original English measure and its German adaptation were applied to the respective language versions [23]. Comparisons were descriptive and conducted separately within each language.

Longitudinal total-score analysis

For each participant, the time was centered on the first V2.1 measurement while maintaining the actual intervals between assessments. A segmented linear regression with participant-fixed effects was fitted to the seven assessments to account for differences in baseline values [24]. The model estimated the slope of V1.0, the immediate change in the value at the transition, and the change in the slope with V2.1. The slope for V2.1 was calculated as the sum of the slope of V1.0 and the change in the slope. Confidence intervals were obtained by participant-cluster bootstrap, resampling complete participant profiles with replacement [25]. We used 10,000 resamples for the model coefficients and 20,000 for the visit-level means, mean residuals, and pointwise confidence bands. Residuals were defined as observed minus fitted total scores. The ± 1-point range was used as a descriptive clinical reference and not as an equivalence margin.

Bias-calibrated item-level effects

For each participant and item, the expected score at the first V2.1 assessment was predicted using a linear trend based on up to the five most recent V1.0 assessments. Predictions were limited to the valid score range (0–4). To account for each participant's typical prediction error, the same method was tested within the V1.0 data by predicting a later V1.0 score from three to five preceding assessments spanning at least 60 days. The mean difference between observed and predicted V1.0 scores was then used to adjust the corresponding difference at the first V2.1 assessment. Mean adjusted differences and 95% confidence intervals were estimated using 10,000 participant-level bootstrap resamples. The ± 0.25-point range was used as a descriptive reference and did not represent an equivalence margin.

Local effects at response boundaries

To determine whether mean item effects masked changes at specific response levels, each item was analyzed at four cumulative boundaries: 0 vs. 1–4, 0–1 vs. 2–4, 0–2 vs. 3–4, and 0–3 vs. 4. For each boundary, the probability of scoring above the boundary at the first V2.1 assessment was predicted from up to the five most recent V1.0 assessments using a linear probability model. Predictions were limited to the range from 0 to 1 and adjusted for each participant's typical V1.0 prediction error. Effects were averaged across participants and reported as percentage-point differences in the probability of scoring above the respective boundary. Percentile 95% confidence intervals were obtained from 10,000 participant-level bootstrap resamples. Two-sided p values were calculated for each of the 48 comparisons and adjusted for multiple comparisons using the Benjamini–Hochberg procedure [26]. Adjusted p values below 0.05 were considered statistically significant.

Statistical analyses were conducted using R version 4.6.1 (2026-06-26). Statistical results were exported from R and plotted in GraphPad Prism version 11.0.2 (GraphPad Software, Boston, MA, USA).

RESULTS

Study population and cohort selection

Two analysis cohorts were derived from the available longitudinal ALSFRS-R-SE data (Fig. 1). The item-level cohort comprised 420 participants, while the longitudinal total-score cohort included 189 participants. Sociodemographic and clinical characteristics of both cohorts are summarized in Table 1.

Table 1
Sociodemographic and clinical characteristics of the patient cohort.

Parameter	Depiction	Longitudinal total-score analysis, sub cohort (n = 189)	Item-level and response boundary analysis, entire cohort (n = 420)
Sex	female, n (%)	78 (41.3)	172 (41.0)
	male, n (%)	111 (58.7)	248 (59.0)
Age at first V2.1 assessment	years, mean (SD)	62.3 (9.7)	61.8 (10.3)
disease duration at first V2.1 assessment	months, median (IQR)	39.8 (22.8–65.9)	37.4 (23.3–60.7)
Onset phenotype	spinal, n (%)	111 (58.7)	235 (56.0)
	bulbar, n (%)	26 (13.8)	61 (14.5)
	axial, n (%)	1 (0.5)	4 (1.0)
	respiratory, n (%)	0 (0)	4 (1.0)
	missing, n (%)	51 (27.0)	116 (27.6)
ALSFRS-R-SE at first V2.1 assessment	mean, (SD)	33.2 (8.4)	31.5 (9.8)
Progression rate V1.0, loss of points/month	median (IQR)	−0.40 (− 0.65 – −0.12)	−0.41 (− 0.81 – −0.13)
V1.0 assessments per participant	median (IQR)	26 (16–39)	16 (10–30)
V2.1 assessments per participant	median (IQR)	4 (3–4)	3 (2–4)
Pre-switch observational period	months, median (IQR)	13.7 (7.7–22.7)	12.4 (6.6–21.7)
Interval between last V1.0 and first V2.1	days, median (IQR)	15.7 (14.2–19.5)	20.9 (15.1–42.8)

Development of ALSFRS-R-SE V2.1

After the revision of ALSFRS-R-SE from V1.0 to V2.1, the original scale structure was retained, while the text length was substantially reduced in both the German and English versions (Table 2; please refer to the supplement section for ALSFRS-R-SE V2.1 in German language **Supplement 1** and English language **Supplement 2**). The reductions were particularly pronounced in the introductory text and in the explanations of the response options. In addition to the reduction in text length, V2.1 was characterized by structural simplification, including more concise descriptions of functional states, fewer parenthetical insertions, and greater separation between instructions and response descriptions. Readability was

calculated using the FRE for the overall text and for the individual items (Figure S1 and Table S1 in **Supplement 3**).

Table 2
Summary of textual and linguistic revisions from ALSFRS-R-SE V1.0 to V2.1 in the German and English versions. The component “Additional item text” comprises item headings, notes, and instructions. The FRE was calculated from pooled introductory text, item-specific notes and response descriptions.

Component	Language	V1.0	V2.1	Change
Introduction (words)	German	158	86	−45.6%
	English	193	70	−63.7%
Response labels (words)	German	357	357	0.0%
	English	332	375	+ 13.0%
Explanations of response options (words)	German	1,225	952	−22.3%
	English	1,315	921	−30.0%
Parenthetical insertions	German	21	4	−81.0%
	English	17	1	−94.7%
Additional item text (words)	German	192	228	+ 18.8%
	English	217	206	−5.1%
Overall word count	German	1,932	1,623	−16.0%
	English	2,057	1,572	−23.6%
Flesch Reading Ease (FRE points)	German	40.8	47.3	+ 6.4
	English	41.3	65.3	+ 23.9

Longitudinal continuity across the version transition

Segmented longitudinal analysis showed a continuous decline in total ALSFRS-R-SE scores across all consecutive self-assessments **(**Fig. 2**)**. The estimated progression rate prior to the transition was − 0.570 points per month (95% bootstrap CI − 0.727 to − 0.414). With transition to V2.1, the estimated immediate change in total score was small (− 0.062 points, 95% confidence interval − 0.306 to 0.184). The estimated change in the slope after the transition was + 0.186 points per month (95% CI − 0.054 to 0.420), resulting in a post-transition decline of − 0.384 points per month (95% confidence interval − 0.615 to − 0.169). The model residuals remained centered near zero throughout the transition, with no apparent systematic deviation associated with the introduction of V2.1.

Item-level continuity across the version transition

In the item-level cohort of 420 participants, the progression-adjusted differences between V1.0 and V2.1 were small for all 12 ALSFRS-R-SE items (Fig. 3A) and ranged from − 0.040 to + 0.036 item points. All 95% confidence intervals for the 12 item-level effects included the value zero. After correcting for multiple comparisons, none of the 48 effects at the response boundaries remained statistically significant. The estimated effects were predominantly small (Fig. 3B). The most pronounced localized shift occurred at the upper boundary of Item 7 (turning over in bed and adjusting bed clothes), where the probability of assessing 4 instead of 0–3 was 4.4 percentage points lower in V2.1 than expected based on the previous trajectory of V1.0 (95% CI − 7.10 to − 1.69 percentage points). None of the 48 probability effects remained significant after false discovery rate correction.

DISCUSSION

Although the ALSFRS-R-SE V1.0 was already explicitly designed to be an easily applicable instrument, its revision has further reduced textual burden and simplified its linguistic arrangement. The analysis indicates that Version 2.1 was integrated into longitudinal remote digital self-assessment without a detectable systematic discontinuity regarding both the total score and individual item scores.

A version-related measurement discontinuity may manifest as an immediate change in total score level or a change in the subsequent longitudinal slope. The estimated immediate total score change was small, and its confidence interval included zero, indicating no detectable abrupt shift in score at the observed V1.0-V2.1 transition. This finding supports longitudinal continuity across the version transition under the conditions examined in this study. The estimated progression rate was slower after the transition, but the change in slope was uncertain and its 95% confidence interval included zero. The observed difference may reflect natural longitudinal variability [27], the relatively short post-transition observation period, or other time-related factors and cannot be attributed specifically to the version change. The item-level analyses were consistent with the total-score findings: all 12 bias-calibrated item-effect confidence intervals included zero, and none of the 48 response-boundary effects remained statistically significant after multiple-comparison correction. The largest numerical deviation at the boundary occurred at the upper limit of Item 7, but there was no corresponding effect for the whole item. This locally limited finding may represent chance variation, particularly because it did not remain statistically significant after correction for multiple comparisons and was not accompanied by a corresponding whole-item effect. It therefore does not provide sufficient evidence of a systematic shift in this response boundary or broader recalibration of the item.

The practical benefit of this consistency is that it may allow further refinement of a widely used instrument while preserving the interpretability of existing longitudinal datasets. In ALS, where patients continue to perform self-assessment despite progressive motor impairment over the course of the disease, removing barriers and thereby reducing avoidable text-related burdens is in itself an important objective [12]. Improvements in readability, particularly through reduced complexity and text shortening, have led to a considerable improvement in the Flesch Reading Ease score. Although usability was not directly evaluated, the more concise and consistent wording of V2.1 aligns with current

recommendations for the electronic implementation of patient-reported outcome measures [15], while readability-guided simplification of patient-facing medical information has been associated with high comprehension and improved topic-specific health literacy [28].

Previous studies on the ALSFRS-R have demonstrated the feasibility of remote digital assessment, as well as the need for standardization with regard to various modes of administration [8, 9, 29]. The 2025 FDA guidelines also emphasize that changes to established clinical outcome assessments, including changes to the wording or response options, should be documented and supported by evidence that is proportionate to the potential impact on the interpretation of the results [30]. By combining a documented revision process with analyses at the total-score, item, and response-boundary levels, this study provides empirical evidence relevant to evaluating the potential measurement impact of the revision and offers a framework for future adjustments that is consistent with the call for international harmonization of the self-administered ALSFRS-R versions [18].

This study has several limitations. The analyses were exploratory and were based neither on a predefined hypothesis nor on a formal equivalence margin. The absence of statistically significant discontinuities therefore suggests consistency but does not prove formal equivalence. Because V1.0 and V2.1 were administered sequentially rather than concurrently, potential version effects could not be completely separated from natural longitudinal changes occurring over time. The requirement to submit repeated ratings of the ALS App before and after the version switch led to the selection of digitally engaged and compliant participants and may have resulted in the underrepresentation of individuals of certain age groups, with severe disabilities, or lower digital competence. The degree of this selection bias cannot be quantified, as no information was available regarding the reasons for missing assessments or the discontinuation of app use. Finally, the study evaluated measurement continuity in a remote digital setting but did not directly examine usability, comprehensibility, or cross-linguistic equivalence.

The ALSFRS-R-SE V2.1 showed no detectable systematic discontinuity across the observed transition from V1.0 to V2.1 while providing more concise explanations and instructions for remote digital self-assessment. Future linguistic and cultural adaptations should follow comparable principles of structured revision and evaluation of measurement continuity. Such a coordinated, multilingual European framework could facilitate the future evaluation and harmonization of longitudinal data across countries and proactively contribute to the international harmonization of the ALSFRS-R-SE remote digital self-assessment for clinical research and patient care and management.

Declarations

Competing Interests

T.M. and C.M. are founders and shareholders of the Ambulanzpartner Soziotechnologie APST GmbH, which makes the mobile application “ALS App”. All other authors have no competing interests to declare that are relevant to the content of this article.

Author Contribution

All authors contributed to the study conception and design. Material preparation and linguistic editing of the questionnaire were performed by A.M., M.B., P.R., T.G, J.v.d.G., P.L., S.S., and T.M. A.M., B.W., S.Sp., P.S., S.S., A.R., C.M., and T.M. were responisble for data collection and analyzed the data. The first draft of the manuscript was written by A.M. and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Acknowledgement

The authors thank all participating patients for making their data available for research and the Boris Canessa ALS Stiftung (Düsseldorf, Germany) for providing continued support.

Data Availability

The datasets supporting the findings of this study were generated for scientific purposes and are held by Ambulanzpartner Soziotechnologie APST GmbH. Restrictions apply to the availability of these data, which are not publicly available. The data may, however, be made available upon reasonable request (s.spittel@ambulanzpartner.de).

References

1. Cedarbaum, J. M., Stambler, N., Malta, E., Fuller, C., Hilt, D., Thurmond, B., & Nakanishi, A. (1999). The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function. *Journal of the Neurological Sciences*, *169*, 13–21. https://doi.org/10.1016/S0022-510X(99)00210-5
2. Kaufmann, P., Levy, G., Montes, J., Buchsbaum, R., Barsdorf, A. I., Battista, V., Arbing, R., Gordon, P. H., Mitsumoto, H., Levin, B., Thompson, J. L. P., Kaufmann, P., Levy, G., Montes, J., Buchsbaum, R., Barsdorf, A. I., Battista, V., Arbing, R., Gordon, P. H., Mitsumoto, H., Levin, B., & Thompson, J. L. P. (2007). Excellent inter-rater, intra-rater, and telephone-administered reliability of the ALSFRS-R in a multicenter clinical trial. *Amyotrophic Lateral Sclerosis*, *8*, 42–46. https://doi.org/10.1080/17482960600888156
3. Shefner, J. M., Bunte, T., Kittle, G., Genge, A., & Van Den Berg, L. H. (2024). Harmonized standard operating procedures for administering the ALS functional rating scale-revised. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *25*, 26–33. https://doi.org/10.1080/21678421.2023.2260832
4. Bakers, J. N. E., de Jongh, A. D., Bunte, T. M., Kendall, L., Han, S. S., Epstein, N., Lavrov, A., Beelen, A., Visser-Meily, J. M. A., van den Berg, L. H., & van Eijk, R. P. A. (2022). Using the ALSFRS-R in multicentre clinical trials for amyotrophic lateral sclerosis: potential limitations in current standard

operating procedures. *Amyotroph Lateral Scler Frontotemporal Degener*, *23*, 500–507. https://doi.org/10.1080/21678421.2021.2016838

5. Hamilton, J., Mohan, P., Kittle, G., & Shefner, J. M. (2023). Impact of mode of training and recertification on ALSFRS-R rater performance. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *24*, 289–294. https://doi.org/10.1080/21678421.2022.2149344
6. Bakker, L. A., Schröder, C. D., Tan, H. H. G., Vugts, S. M. A. G., van Eijk, R. P. A., van Es, M. A., Visser-Meily, J. M. A., & van den Berg, L. H. (2020). Development and assessment of the inter-rater and intra-rater reproducibility of a self-administration version of the ALSFRS-R. *Journal Of Neurology, Neurosurgery And Psychiatry*, *91*, 75–81. https://doi.org/10.1136/jnnp-2019-321138
7. Montes, J., Levy, G., Albert, S., Kaufmann, P., Buchsbaum, R., Gordon, P. H., & Mitsumoto, H. (2006). Development and evaluation of a self-administered version of the ALSFRS-R. *Neurology*, *67*, 1294–1296. https://doi.org/10.1212/01.wnl.0000238505.22066.fc
8. Maier, A., Holm, T., Wicks, P., Steinfurth, L., Linke, P., Münch, C., Meyer, R., & Meyer, T. (2012). Online assessment of ALS functional rating scale compares well to in-clinic evaluation: A prospective trial. *Amyotrophic Lateral Sclerosis*, *13*, 210–216. https://doi.org/10.3109/17482968.2011.633268
9. Meyer, T., Spittel, S., Grehl, T., Weyen, U., Steinbach, R., Kettemann, D., Petri, S., Weydt, P., Günther, R., Baum, P., Schlapakow, E., Koch, J. C., Boentert, M., Wolf, J., Grosskreutz, J., Rödiger, A., Ilse, B., Metelmann, M., Norden, J., Koc, R. Y., Körtvélyessy, P., Riitano, A., Walter, B., Hildebrandt, B., Schaudinn, F., Münch, C., & Maier, A. (2022). Remote digital assessment of amyotrophic lateral sclerosis functional rating scale – a multicenter observational study. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, 1–10. https://doi.org/10.1080/21678421.2022.2104649
10. Maier, A., Koc, Y., Steinfurth, L., Kettemann, D., Norden, J., Riitano, A., Schmitt, P., Kolzarek, F., Subramanian, S., Münch, C., Spittel, S., & Meyer, T. (2026). Agreement Between the Harmonized and the Self-Explanatory Versions of the Revised ALS Functional Rating Scale in a Clinical Setting. *Muscle and Nerve*, *73*, 250–259. https://doi.org/10.1002/mus.70092
11. Van Eijk, R. P. A., Beelen, A., Kruitwagen, E. T., Murray, D., Radakovic, R., Hobson, E., Knox, L., Helleman, J., Burke, T., Rubio Pérez, M. Á., Reviers, E., Genge, A., Steyn, F. J., Ngo, S., Eaglesham, J., Roes, K. C. B., Van Den Berg, L. H., Hardiman, O., & McDermott, C. J. (2021). A Road Map for Remote Digital Health Technology for Motor Neuron Disease. *Journal Of Medical Internet Research*, *23*, e28766. https://doi.org/10.2196/28766
12. Aiyegbusi, O. L., Roydhouse, J., Rivera, S. C., Kamudoni, P., Schache, P., Wilson, R., Stephens, R., & Calvert, M. (2022). Key considerations to reduce or address respondent burden in patient-reported outcome (PRO) data collection. *Nature Communications*, *13*, 6026. https://doi.org/10.1038/s41467-022-33826-4
13. Maier, A., Boentert, M., Reilich, P., Witzel, S., Petri, S., Großkreutz, J., Metelmann, M., Lingor, P., Cordts, I., Dorst, J., Zeller, D., Günther, R., Hagenacker, T., Grehl, T., Spittel, S., Schuster, J., Ludolph, A., Meyer, T., & for the MND-NET consensus group. (2022). ALSFRS-R-SE: an adapted, annotated, and self-

explanatory version of the revised amyotrophic lateral sclerosis functional rating scale. *Neurol Res Pract*, *4*, 60. https://doi.org/10.1186/s42466-022-00224-6

14. Steinfurth, L., Grehl, T., Weyen, U., Kettemann, D., Steinbach, R., Rödiger, A., Grosskreutz, J., Petri, S., Boentert, M., Weydt, P., Bernsen, S., Walter, B., GüNTHER, R., Lingor, P., Koch, J. C., Baum, P., Weishaupt, J. H., Dorst, J., Koc, Y., Cordts, I., Vidovic, M., Norden, J., Schumann, P., Körtvélyessy, P., Spittel, S., Münch, C., Maier, A., & Meyer, T. (2025). Self-assessment of amyotrophic lateral sclerosis functional rating scale on the patient's smartphone proves to be non-inferior to clinic data capture. Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration 1–12. https://doi.org/10.1080/21678421.2025.2468404
15. Mowlem, F. D., Elash, C. A., Dumais, K. M., Haenel, E., O'Donohoe, P., Olt, J., Kalpadakis-Smith, A. V., James, B., Balestrieri, G., Becker, K., Newara, M. C., & Kern, S. (2024). Best Practices for the Electronic Implementation and Migration of Patient-Reported Outcome Measures. *Value in Health*, *27*, 79–94. https://doi.org/10.1016/j.jval.2023.10.007
16. Curtis, S. E., Tatlock, S., O'Hara, L., Seçinti, E., Mehdiyoun, N. F., Ayala-Nunes, L., Flynn, J., Fernelius, K., Hodson, N., & Delbecque, L. (2026). Patient experience and clinical outcome assessment validity in amyotrophic lateral sclerosis: a targeted literature review. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *27*, 285–300. https://doi.org/10.1080/21678421.2025.2604233
17. Wild, D., Grove, A., Martin, M., Eremenco, S., McElroy, S., Verjee-Lorenz, A., & Erikson, P. (2005). Principles of Good Practice for the Translation and Cultural Adaptation Process for Patient-Reported Outcomes (PRO) Measures: Report of the ISPOR Task Force for Translation and Cultural Adaptation. *Value in Health*, *8*, 94–104. https://doi.org/10.1111/j.1524-4733.2005.04054.x
18. Allen, M. D., Van Eijk, R. P. A., Knox, L., Carlton, J., Hobson, E., Mcdermott, C. J., Murray, D., Berry, J., Meyer, T., & Genge, A. (2025). Variability across versions of the self-administered ALSFRS-R: a review and call for harmonization. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *1–6*. https://doi.org/10.1080/21678421.2025.2522400
19. Genge, A., Cedarbaum, J. M., Shefner, J., Chio, A., Al-Chalabi, A., Van Damme, P., McDermott, C., Glass, J., Berry, J., Van Eijk, R. P. A., Fournier, C., Grosskreutz, J., Andrews, J., Bertone, V., Bunte, T. M., Couillard, M., Cummings, C., Kittle, G., Polzer, J., Salmon, K., Straub, C., & Van Den Berg, L. H. (2024). The ALSFRS-R Summit: a global call to action on the use of the ALSFRS-R in ALS clinical trials. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *25*, 382–387. https://doi.org/10.1080/21678421.2024.2320880
20. Von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., & Vandenbroucke, J. P. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, *61*, 344–349. https://doi.org/10.1016/j.jclinepi.2007.11.008
21. Shefner, J. M., Al-Chalabi, A., Baker, M. R., Cui, L-Y., De Carvalho, M., Eisen, A., Grosskreutz, J., Hardiman, O., Henderson, R., Matamala, J. M., Mitsumoto, H., Paulus, W., Simon, N., Swash, M., Talbot, K., Turner, M. R., Ugawa, Y., Van Den Berg, L. H., Verdugo, R., Vucic, S., Kaji, R., Burke, D., &

Kiernan, M. C. (2020). A proposal for new diagnostic criteria for ALS. *Clinical Neurophysiology*, *131*, 1975–1978. https://doi.org/10.1016/j.clinph.2020.04.005

22. Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology*, *32*, 221–233. https://doi.org/10.1037/h0057532
23. Betschart, P., Staubli, S. E., Zumstein, V., Babst, C., Sauter, R., Schmid, H-P., & Abt, D. (2019). Improving Patient Education Materials: A Practical Algorithm from Development to Validation. *Current Urology*, *13*, 64–69. https://doi.org/10.1159/000499291
24. Lopez Bernal, J., Cummins, S., & Gasparrini, A. (2016). Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International Journal Of Epidemiology*. https://doi.org/10.1093/ije/dyw098. dyw098.
25. Field, C. A., & Welsh, A. H. (2007). Bootstrapping Clustered Data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *69*, 369–390. https://doi.org/10.1111/j.1467-9868.2007.00593.x
26. Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x
27. Van Eijk, R. P. A., Weemering, D. N., Opie-Martin, S., Van Unnik, J. W. J., Caravaca Puchades, A., Chiò, A., Corcia, P., Galvin, M., Hardiman, O., Heverin, M., Hobin, F., Holmdahl, O., Ingre, C., Lamaire, N., Mac Domhnaill, É., McDonough, H., Manera, U., McDermott, C. J., McFarlane, R., Mouzouri, M., Ombelet, F., Povedano Panadés, M., Sennfält, S., Shaw, P. J., Terrafeta Pastor, C., Van Damme, P., Vasta, R., Veldink, J. H., Al-Chalabi, A., & Van Den Berg, L. H. (2025). Natural history of the revised ALS functional rating scale and its association with survival: the PRECISION-ALS Extant Study. *Amyotrophic Lateral Sclerosis and Frontotemporal Degeneration*, *26*, 30–40. https://doi.org/10.1080/21678421.2024.2443985
28. Amato, R., Del Toro-Pagan, N. M., Nguyen, H., Plummer, J., Pizzolato, K., Krause, D., & Dowd, D. (2025). Assessing the Impact of Simplified Language on a Patient-Facing Pharmacogenetic Report: A User Comprehension Study. *JPM*, *15*, 247. https://doi.org/10.3390/jpm15060247
29. Bhushan, N. L., Romano, C. D., Gras-Najjar, J., Reno, J., Rockwood, N., Quattrone, W., Adams, E. T., Kelly, B., McLeod, L., Bhavnani, S. P., Bocell, F. D., Campbell, M., Kontson, K., Reasner, D., Zhang, C., & Retzky, S. (2024). Remote-Use Applications of the Amyotrophic Lateral Sclerosis Functional Rating Scale-Revised Clinical Outcome Assessment Tool: A Scoping Review. *Value in Health*, *27*, 1454–1465. https://doi.org/10.1016/j.jval.2024.05.005
30. Food, U. S. (2025). and Drug Administration Patient-Focused Drug Development: Selecting, Developing, or Modifying Fit-for-Purpose Clinical Outcome Assessments. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-focused-drug-development-selecting-developing-or-modifying-fit-purpose-clinical-outcome Accessed 25 September 2026.

Figures

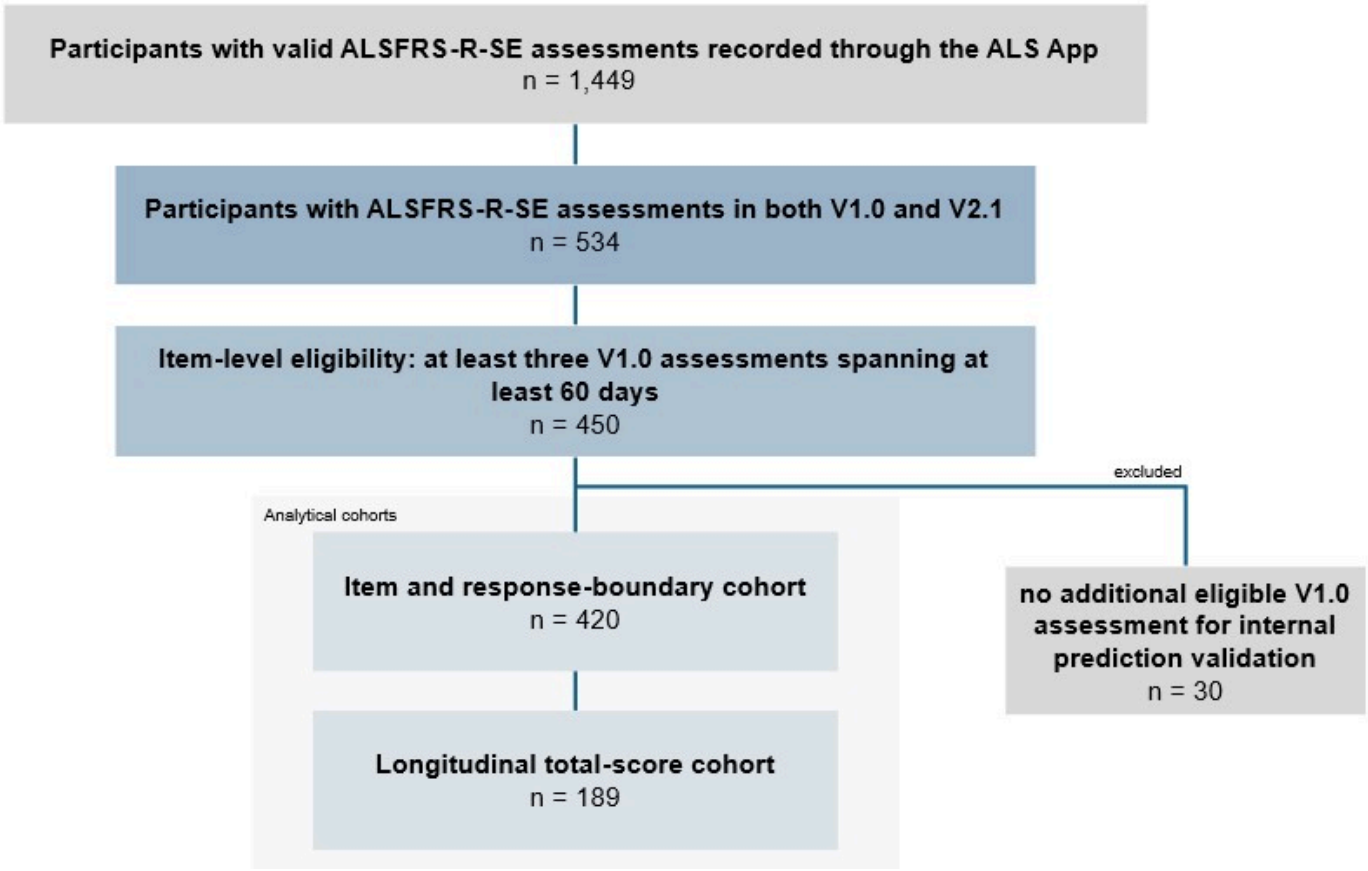


Figure 1

Flow chart of participant selection and derivation of the analytical cohorts.

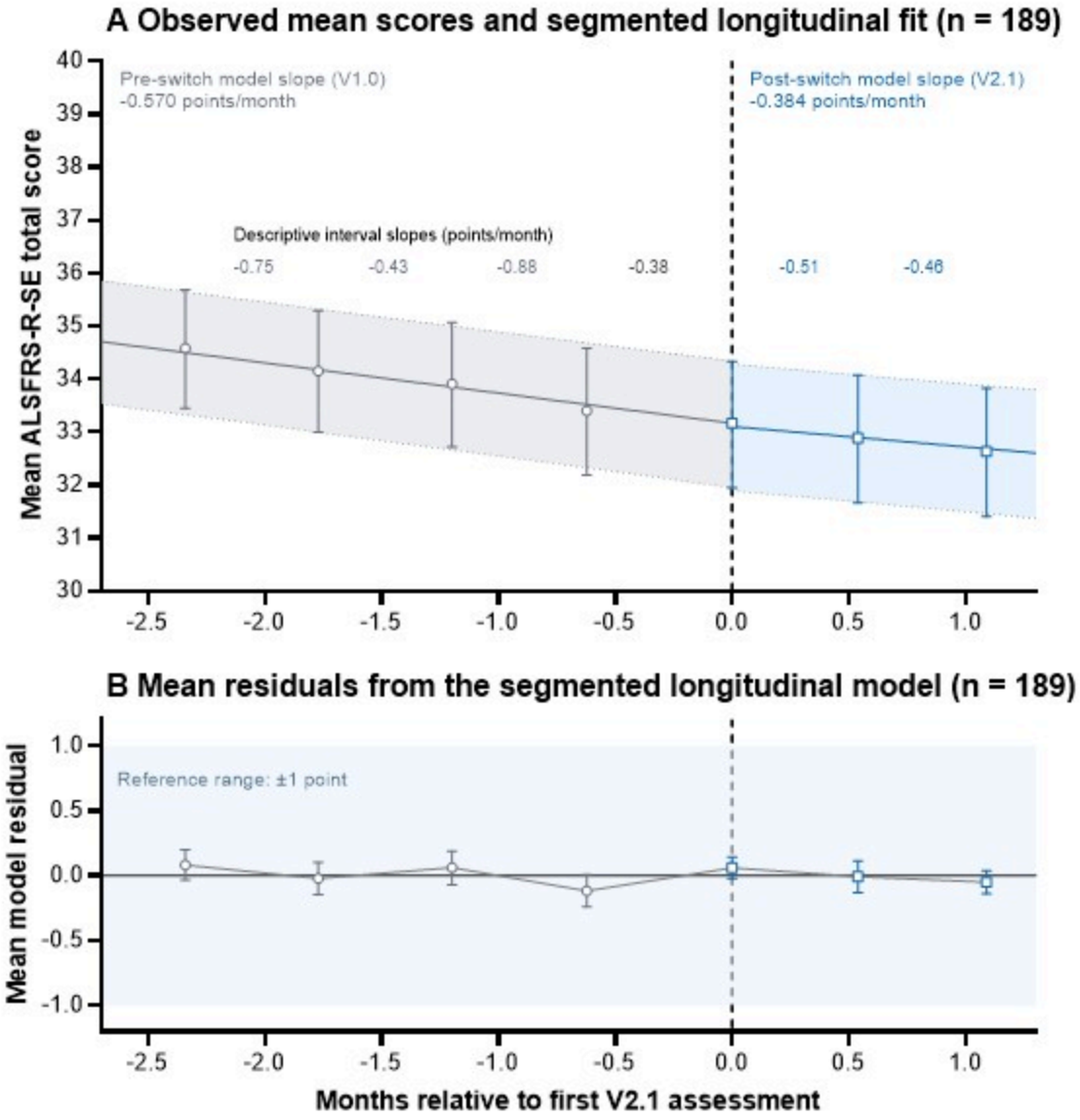


Figure 2

Longitudinal continuity of ALSFRS-R-SE total scores across the transition from V1.0 to V2.1. Data are included from 189 participants with four V1.0 and three V2.1 assessments. The timeline is depicted relative to the first V2.1 assessment (month 0; vertical dashed line). **(A)** Observed mean ALSFRS-R-SE total scores (V1.0, gray circles; V2.1, blue squares) and 95% confidence intervals. Solid lines and shaded bands show estimates and 95% confidence intervals from a segmented longitudinal model allowing total score and slope changes at the transition. Values between assessments are descriptive slopes based on observed means. **(B)**Mean observed-minus-predicted residuals with 95% confidence intervals, zero indicates perfect model agreement. The reference range of ±1-point is descriptive.

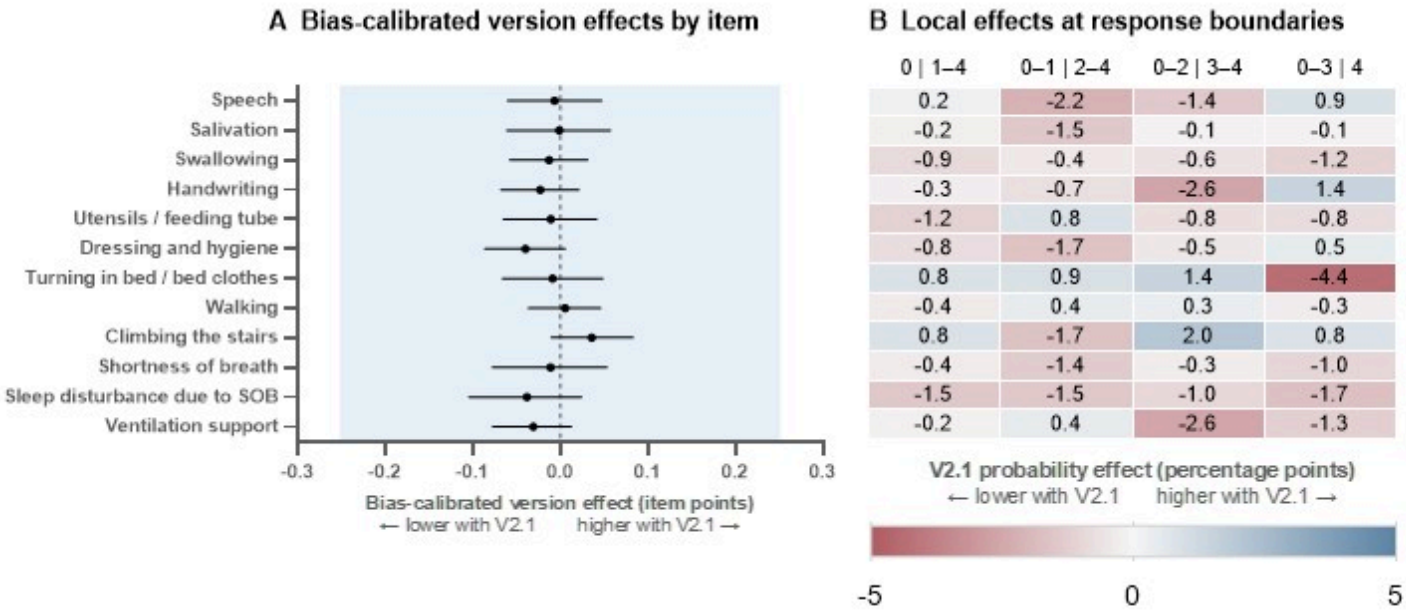


Figure 3

Item-level continuity and exploratory response-boundary effects across the ALSFRS-R-SE V1.0-to-V2.1 transition. (A) Bias-calibrated V2.1-minus-V1.0 item effects are shown with patient-cluster bootstrap 95% confidence intervals. The dashed line indicates no effect; the light-blue area represents a descriptive ±0.25-point range. **(B)**V2.1-minus-V1.0 differences in the estimated probability of scoring above each cumulative response boundary are expressed in percentage points. Column labels indicate the score ranges below and above each boundary. Cell values and colors represent the magnitude and direction of these differences: red indicates a lower probability under V2.1, blue a higher probability, and white zero difference. None of the 48 direct probability effects remained significant after false discovery rate correction. SOB = shortness of breath.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplement1.pdf
- Supplement2.pdf
- Supplement3.pdf