The original Amyotrophic Lateral Sclerosis Functional Rating Scale (ALSFRS), developed in the 1990s, was designed as a clinical tool to help with the clinical assessment of patients with ALS.44,45 Prior to the development of the original version of the ALSFRS, several instruments for measuring disease status and progression in ALS were used in clinical trials and practice, such as the Norris scale, Baylor scale, and Tufts Quantitative Neuromuscular Examination. However, these metrics failed to consider all areas of functionality that may be impacted by ALS.25 The developers of the ALSFRS sought to create a scale that was easy to use and assessed multiple clinically relevant factors, including a patient’s ability to conduct activities of daily living (ADLs).25 The original ALSFRS, like the later revised version, was a questionnaire-based scale that measured physical function in carrying out ADLs of patients with ALS.25 It posed a series of questions regarding a patient’s ability to carry out functions in 4 domains: gross motor tasks, fine motor tasks, bulbar function, and respiratory function.25

The ALSFRS was validated and employed in both clinical trials and practice; however, over time, it became evident that respiratory function was not weighted in the overall assessment as well as the other 3 domains were.25 Therefore, a revised ALSFRS (ALSFRS-R) was developed in 1999 to include additional considerations relevant to a patient’s respiratory function.25 It consists of 12 questions, 3 for each of the 4 domains, and the scoring for each question ranges from 0 (no ability) to 4 (normal ability), producing a potential total maximum score of 48 (Table 2).25

The ALSFRS-R is now established not only for clinical use to measure disease progression and patient status, but also as a standard outcome measure employed across ALS trials.1,17 The ALSFRS-R provides a uniform measure of efficacy that is easily administered in person or remotely, which may improve patient retention since it minimizes the number of clinic visits.24 It has a high inter-rater and intra-rater reliability, and scores are strongly correlated with quality of life.2,24,25

Although the ALSFRS-R encompasses many clinically relevant features of disease progression, it has limitations.1 It provides a uniform efficacy measurement, but the scoring metric remains subjective.2 It may also have a potential floor effect in advanced disease. Further, the questionnaire includes inquiries regarding symptoms that can be potentially resolved by supportive treatment; thus, it is difficult to compare patients’ scores since some symptoms may be masked by therapy.2 The ALSFRS-R has also been challenged...
on the basis of inconsistency between steps of change (ie, point scoring) within each domain, as it is a nominal scale. Support for this contention was provided in a study showing that the motor subscore deteriorated at earlier time point in spinal-onset disease compared with bulbar-onset, and the bulbar subscore deteriorated both earlier and more rapidly in bulbar-onset disease.\textsuperscript{44,48} The risk of specific features of change being concealed due to inconsistent equivalence between domains of the ALSFRS-R may be more intrinsic in a disease with heterogeneous patterns of progression like ALS.\textsuperscript{44} Notably, Franchignoni and colleagues have demonstrated evidence of multidimensionality in the ALSFRS-R, suggesting that subscores presented separately rather than combined as 1 single score might better reflect patient status.\textsuperscript{48,49}

While the ALSFRS-R score is a well-validated scoring metric used in clinical trials, its limitations should be considered when evaluating clinical trial data.

APPENDIX II. CLINICAL TRIALS OF EDARAVONE IN THE TREATMENT OF ALS

Mitsubishi Tanabe Pharma America sponsored an edaravone clinical development program with the goal of both bringing effective treatments to patients with ALS and uncovering methods to develop more efficient clinical trials to advance the treatment of ALS and globally aid future clinical trials.\textsuperscript{7} Spanning 13-plus years, the edaravone clinical development program was a multi-study program involving a series of randomized clinical trials, extension trials, and post hoc analysis studies. Post hoc analyses were conducted to determine the limitations of a given trial and how to improve trial designs of subsequent trials. The details of each trial are discussed below.

Study 16 (MCI 186-16): The First Phase 3 Trial of Edaravone in ALS

Methods

Study 16 was a randomized, double-blind, parallel-group, placebo-controlled phase 3 trial conducted at 29 sites in Japan from May 2006 to September 2008.\textsuperscript{28} It was a 2-phase, 36-week study, including a 12-week pre-observation period and a 24-week treatment period.\textsuperscript{28}

Patients were eligible if they met the following inclusion criteria:\textsuperscript{28}

- Aged 20 to 75 years
- Diagnosis of “definite,” “probable,” or “laboratory-supported-probable” ALS
- % predicted forced vital capacity (%FVC) $\geq$70%
- Disease duration $\leq$3 years
- Change in ALSFRS-R score during the pre-observational period of $-1$ to $-4$ points
- Japanese ALS severity classification of 1 or 2

Patients were excluded if any of the following were met:\textsuperscript{28}

- Deteriorated general condition as judged by investigators
- Reduced respiratory function or dyspnea
- Creatinine clearance of $\leq$50 ml/min within 28 days of study treatment initiation
- Complications that might substantially influence evaluation of drug efficacy or require hospitalization
- Infection requiring antibiotic therapy
- Undergoing cancer treatment

Patients were randomized to receive edaravone 60 mg once-daily intravenous (IV) infusion or placebo for the 24-week treatment period.\textsuperscript{28} There were 6 total treatment cycles, where each cycle is 4 weeks (cycles 1 to 6). The first cycle consisted of 14 days of drug administration followed by 14 days off drug. All subsequent cycles (cycles 2-6) were 10 of 14 days on drug, followed by 14 days off drug.\textsuperscript{28}

The primary efficacy end point was change in ALSFRS-R score. Secondary end points included changes in %FVC, grip strength, pinch strength, Modified Norris Scale score, 40-item ALS Assessment Questionnaire (ALSAQ-40), and time to death or specified state of disease progression. Evaluations were conducted prior to the pre-observation period, prior to the start of the first treatment cycle, and at the end of each treatment cycle after the 14-day observation period.\textsuperscript{28} Safety outcomes were assessed in terms of number and severity of adverse events (AEs), adverse drug reactions, and the results of clinical laboratory tests and sensory tests.\textsuperscript{28}

Results

Baseline characteristics. A total of 206 patients were randomized; however, 1 was excluded due to misdiagnosis. Of the 205 patients who were studied, 101 patients were treated with edaravone and 104 patients received placebo.\textsuperscript{28} Baseline characteristics between the 2 study groups were fairly similar. Approximately 64% of the patients were male, and median age was about 58 years. The vast majority of the patients initially presented with limb-onset disease (81%-82% versus 18%-19% bulbar-onset), and median duration of
EDARAVONE IN THE TREATMENT OF ALS

Efficacy. The adjusted mean change in ALSFRS-R score from baseline to the end of treatment period was –5.70 ± 0.85 in the edaravone group and –6.35 ± 0.84 in the placebo group. Although there was less of a decline in ALSFRS-R score seen in the edaravone group than the placebo group, no statistically significant difference was found (P = .411).

Safety. There was no significant difference in AEs reported between the edaravone group and placebo group. The most common AEs reported in the edaravone and placebo groups, respectively, were nasopharyngitis (21.6% vs 21.2%), gait disturbance (19.6% vs 15.4%), constipation (12.7% vs 16.3%), dysphagia (7.8% vs 11.3%), and contusion (11.8% vs 4.8%). A total of 17.6% of patients in the edaravone group versus 23.1% in the placebo group reported experiencing serious AEs (SAEs), which included dysphagia (7.8% vs 10.6%), gait disturbance (2.9% vs 1.9%), and muscular weakness (1% in each group). A total of 5 deaths, secondary to respiratory dysfunction, were reported (3 in the edaravone group and 2 in the placebo group); however, the investigators attributed the deaths to the underlying disease rather than the study drug. No serious adverse drug reactions (ADRs) occurred in either group.

Study 16 Post Hoc Analysis

Study 16 was the first phase 3 clinical trial of the edaravone clinical development program. Although there was less of a decline in functionality (measured by the ALSFRS-R score) seen in the edaravone group than the placebo group, the difference was not statistically significant. Given the large unmet need for more treatment options for ALS, the investigators sought to continue the development of edaravone by discovering more efficient trial design strategies. The investigators conducted an in-depth post hoc subgroup analysis of Study 16 to determine whether the failure to find a significant difference in change in ALSFRS-R score was due to limitations of the Study 16 trial design, and to determine how they could address these limitations in future trials.

Methods

Hypothsis for identifying the subpopulation. The investigators hypothesized that a significant difference in change in ALSFRS-R score had not been found in Study 16 because the clinical courses of the patients enrolled had such large variations. The large variations resulted in an exceedingly wide range of changes in the ALSFRS-R score (Figure 2). Given the limited treatment duration of the clinical trial and low prevalence of patients with ALS, the investigators set out to identify a patient subpopulation in whom the changes in the ALSFRS-R score in a 24-week period would be
The investigators identified the following subpopulations of patients with ALS: more homogenous subpopulation with ALS would increase the statistical power without needing to increase the trial duration. Extending the duration of a trial in which a placebo-controlled group is utilized would raise ethical concerns regarding allowing patients to remain untreated on placebo.

Subpopulations. The investigators identified the following subgroups of patients with ALS in Study 16:

- **Step 1:** Efficacy-Expected Subpopulation (EESP): baseline characteristics of %FVC ≥80% and ≥2 points for all item scores in the ALSFRS-R.
- **Step 2:** Greater-Efficacy-Expected Subpopulation (dpEESP2y): same baseline characteristics of EESP (%FVC ≥80% and ≥2 points for all ALSFRS-R items) plus diagnosis of ‘definite’ or ‘probable’ ALS and within 2 years of initial ALS symptom onset.

The investigators explained the rationale for these subpopulation criteria. They excluded patients with advanced disease because their baseline ALSFRS-R score on some items may have already declined to 0 or 1. Patients with advanced disease could have no decline or a maximum decline of 1 point in any given item; this could potentially and inappropriately bias the results. Thus, the criterion requiring a score of ≥2 points for each item of the ALSFRS-R was instituted. The investigators determined that patients with respiratory dysfunction would also inappropriately skew the efficacy data since some patients with respiratory dysfunction may show particularly rapid progression, thereby masking the potential efficacy of edaravone. Therefore, the criterion requiring %FVC ≥80% was added. The investigators eliminated the inclusion of patients with “probable-laboratory-supported” diagnosis in order to focus on patients with a diagnosis of “definite” or “probable.”

Changes in ALSFRS-R scores within the 24-week treatment period in patients with a “probable-laboratory-supported” diagnosis were smaller and could thus potentially bias the results. Based on ALS expert opinion, the investigators determined that the inclusion of patients who were within 2 years of ALS symptom onset would be more appropriate for a clinical trial with a 24-week treatment period. The rationale was that patients with more advanced ALS might exhibit less pronounced changes in ALSFRS-R scores.

The investigators evaluated the Study 16 data in the following subgroups: The EESP subgroup included 104 patients, 54 treated with edaravone and 50 with placebo, while the dpEESP2y subgroup included 72 patients, 40 treated with edaravone and 32 with placebo.

### Results

**Efficacy.** A significantly lower decline in ALSFRS-R score was seen in patients receiving edaravone compared with those receiving placebo in both the EESP and dpEESP2y groups. In the EESP group, the difference in the mean ALSFRS-R score change between the edaravone and placebo arms was 2.20 (±1.03) points (P = .0360). The difference was even greater in the dpEESP2y group, with a 3.01 (±1.33) point difference (P = .0270). Figure 3 compares the intergroup differences in change in ALSFRS-R of the full analysis set (FAS), which is the original Study 16 population, including the EESP and dpEESP2y groups.

**Safety.** The proportions of AEs and SAEs were similar in all randomized patients. However, there was a significantly lower rate of SAEs in patients receiving edaravone versus placebo in both the EESP and dpEESP2y groups. In the EESP group, 1.9% of the edaravone group compared with 4.0% of the placebo group experienced SAEs (P = .0085). In the dpEESP2y group, the rate of SAEs in the edaravone group versus the placebo group was 2.5% versus 25.0%, respectively (P = .0007).

### MCI 186-17 (Study 17):

**An Extension Study of Study 16**

To evaluate the long-term efficacy and safety of edaravone in ALS, the investigators conducted an extension study of Study 16. The study included 180 patients who completed treatment in Study 16 and did not meet the exclusion criteria at the time of enrollment into Study 17.

### Methods

Patients who had received edaravone or placebo in Study 16 (treatment period termed cycles 1-6) were divided into 3 treatment groups for Study 17’s 24-week double-blind phase (cycles 7-12). Patients who had received edaravone during Study 16 were randomized to receive either edaravone (E-E group, n = 48) or placebo (E-P group,
The efficacy end points evaluated included ALSFRS-R score, time to death or specified state of disease progression (death, disability of independent ambulation, loss of upper-limb function, tracheotomy, use of respirator, and use of tube feeding), %FVC, Modified Norris Scale score, ALSAQ-40 score, hand grip strength, and pinch strength. A predetermined subgroup efficacy analysis was conducted of Study 17 patients meeting the EESP criteria defined in the Study 16 post hoc analysis, which consisted of 96 patients (27 in the E-E group, 25 in the E-P group, and 44 in the P-E group).

**Results**

**Efficacy.** The mean change in ALSFRS-R score for the E-E group was \( -4.42 \pm 0.69 \) points compared with \( -5.58 \pm 0.74 \) points for the E-P group. Although a lower decline in ALSFRS-R score favoring edaravone was shown (difference of \( 1.16 \pm 0.93 \)), it was not statistically significant. In the analysis of patients in the EESP subgroup, the mean change in ALSFRS-R score for the E-E group was \( -4.01 \pm 0.86 \) points versus \( -5.86 \pm 0.98 \) points in the E-P group. Although the intergroup difference of \( 1.85 \pm 1.14 \) points was greater than the difference seen in the FAS of Study 17, it was not statistically significant. However, this study involved fewer patients than did Study 16, and a controlled comparison with a group that received placebo for the entire duration of Study 16 through Study 17 was not utilized because of the ethical concerns that would have been involved in long-term administration of placebo to patients with ALS.

**Safety.** There was no significant difference between the E-E and E-P groups for incidence of AEs reported. The largest differences that were seen between the E-E and E-P groups, respectively, were in gait disturbance (29.2% vs 20.0%) and respiratory failure (12.5% vs 4.4%). The most common SAEs across all treatment groups were dysphagia and musculoskeletal disorder. There was a significantly higher occurrence of SAEs in the E-E group (52.1%) versus the E-P group (28.9%) (\( P = .0344 \)). However, the majority of the SAEs were considered attributable to the progression of ALS rather than edaravone. No serious ADRs were seen.

**Post Hoc Analysis of Study 17: Inclusion of the dpEESP2y Group**

A post hoc analysis of Study 17 was conducted to evaluate Study 17 data in a subgroup of patients who met the dpEESP2y criteria of Study 16 post hoc analysis. Because the Study 16 post hoc analysis was still ongoing when the design of Study 17 was planned, a dpEESP2y subgroup analysis was not included in the original Study 17 publication. The purpose of this analysis was to determine whether the trend observed in the Study 16 post hoc analysis, with regard to the efficacy seen in the dpEESP2y subgroup, continued in the 36-week extension Study 17.

**Results**

The dpEESP2y subgroup in this post hoc analysis of Study 17 included 67 patients (22 in the E-E group, 16 in the E-P group, and 29 in the P-E group). The mean change in ALSFRS-R score for the E-E group during the double-blind phase (cycles 7-12) was \( -4.22 \pm 1.04 \) points compared with \( -7.02 \pm 1.39 \) points for the E-P group. This difference of \( 2.79 \pm 1.51 \) points (\( P = .0719 \)) was a greater-between-group difference compared with that seen in the FAS group and in the EESP subgroup in the original Study 17. Although this was not statistically significant, likely due to inherent limitations with post hoc analysis for subpopulations without sufficient statistical power, the results of the post hoc analysis showed a beneficial trend favoring edaravone based on the ALSFRS-R score change.

**Study 18: Edaravone Treatment in More Advanced ALS**

Study 18 was an exploratory double-blind, placebo-controlled trial to assess the efficacy and safety of edaravone in patients with more advanced disease. The inclusion criteria required patients to be grade 3 on the Japanese ALS severity classification (requiring assistance for eating, excretion, or ambulation), with a %FVC≥60% at baseline. As an exploratory study, it only included 25 patients (13 receiving edaravone and 12 receiving placebo); it was not sufficiently powered to find a difference.

**Results**

**Efficacy.** As expected, there was no statistically significant between-group difference observed. The mean change in ALSFRS-R score in the edaravone group was \( -6.52 \pm 1.78 \) points and \( -6.00 \pm 1.83 \) points in the placebo group.

**Safety.** Similar rates of AEs were seen in the edaravone and placebo groups, with no single AE occurring in more than 3 patients with the exception of gait disturbance, which occurred in 4 (30.8%) edaravone patients compared with 1 (8.3%) placebo patient. Three patients in the edaravone group and 2 patients in the placebo group experienced SAEs; dysphagia, affecting 2 patients in the edaravone group, was the only SAE observed in more than 1 patient in either group. The SAEs were attributed to the progression of ALS, and no serious ADRs were seen in either group.

**Study 19: Pivotal Phase 3 Trial**

Study 19 was the culmination of the 13-year edaravone clinical development program. Building on the findings from the prior studies,
the investigators utilized an enrichment strategy in Study 19 to increase statistical power by incorporating the revised inclusion criteria developed in the Study 16 post hoc analysis. Although a significant difference in the change in ALSFRS-R scores were seen in Study 16, there are limitations to the interpretability of post hoc analyses; therefore, Study 19 was developed to further understand the effect of edaravone on ALSFRS-R scores.

Methods
Study 19 was a randomized, double-blind, parallel-group, placebo-controlled phase 3 trial conducted at 31 hospitals in Japan. Study 19 had a similar design to that of Study 16, with 2 study phases: a 12-week pre-observational period and a 24-week treatment period. The inclusion criteria for Study 19 were the revised criteria identified in the Study 16 post hoc analysis (dpEESP2y subgroup). Patients were enrolled if they met the following criteria:

- Aged 20 to 75 years
- Diagnosis of “definite” or “probable” ALS
- %FVC ≥80%
- Disease duration ≤2 years
- Score of ≥2 on all 12 items of the ALSFRS-R
- Change in total ALSFRS-R score of –1 to –4 points during the pre-observational phase
- Japanese ALS severity classification of 1 or 2
- Patients were excluded if any of the following were met:
  - Score of ≤3 for the ALSFRS-R items involving dyspnea, orthopnea, or respiratory insufficiency
  - Creatinine clearance ≤50 ml/min
  - History of spinal surgery after ALS onset

Riluzole was permitted if the patient was already being treated with it, as long as there was no change in regimen. New initiation of riluzole was not permitted.

Similar to Study 16, eligible patients were randomized 1:1 to receive edaravone 60 mg or placebo for the 24-week treatment period involving 6 treatment cycles.

The primary end point was the mean change in ALSFRS-R score from baseline to end of treatment period. Secondary end points included changes in %FVC, Modified Norris Scale score, ALSAQ-40 assessment questionnaire, ALS severity classification, grip and pinch strength, and time to death or disease progression. Evaluations were performed before the 12-week pre-observation period, before cycle 1, and at the end of each cycle.

Results
A total of 137 patients were randomized in the treatment period, 69 in the edaravone group and 68 in the placebo group. However, 3 patients (1 in the edaravone group and 2 in the placebo group) were not included for the FAS due to discontinuation before the end of cycle 3. Thus, 134 patients (edaravone n = 68, placebo n = 66) were included in the FAS.

Baseline characteristics. The baseline characteristics among the 2 groups were well matched, with the exception of proportion of male patients (55% of edaravone versus 60% of placebo group) and proportion of patients who were Japanese ALS severity classification grade 2 (68% of edaravone group versus 76% of placebo group). The average age of the patients was approximately 60 years, with an average duration of disease of 1.06 to 1.13 years. A total of 91% of patients were using riluzole at baseline. The mean baseline ALSFRS-R score was 41.9 points in the edaravone group and 41.8 points in the placebo group.

Efficacy. The mean change in ALSFRS-R scores for the edaravone group was −5.01 ± 0.64 compared with −7.50 ± 0.66 in the placebo group. Between the 2 groups, there was a statistically significant least-squares mean difference in ALSFRS-R scores of 2.49 ± 0.76 (95% CI, −0.99 to 3.98; P = .013). This difference represents 33% less functional loss in patients treated with edaravone compared with those receiving placebo over the course of 24 weeks. Utilizing the criteria for evaluating clinical significance set forth by Castrillo-Viguera and colleagues, the investigators concluded that the 33% suppression of ALSFRS-R score was clinically significant. Castrillo-Viguera and colleagues determined that a suppression of ALSFRS-R score by 25% or more is clinically meaningful, based upon a survey of clinicians.

The reduction in quality of life among patients treated with edaravone was significantly less than the reduction in the placebo group as measured by the ALSAQ-40: least-square means difference was −8.79 (P = .0093). In a recent review discussing the use of edaravone in patients with ALS, researchers noted that there was a favorable effect on quality of life due to improvements in the ALSAQ-40 questionnaire. Additionally, researchers recognized that the results of Study 19 lead to the FDA approval of edaravone, further reinforcing the improved benefit of edaravone seen from Study 19.

Safety. The same proportion of patients in both the edaravone and placebo groups reported AEs (84%). SAEs were less common in the edaravone group (16% vs 24%), and dysphagia was the only SAE that occurred in more than 2 patients in either group. Eight patients (12%) in each. No deaths occurred during the study. A total of 3% of the edaravone group patients experienced ADRs (eg, abdominal discomfort, eczema, and abnormal liver function tests) and 7% of the placebo group patients had ADRs (eg, dizziness, constipation, rash, abnormal liver function tests, chondrocalcinosis pyrophosphate, increased serum bilirubin, and increased serum creatine phosphokinase). No serious ADRs were seen in either group.

24-Week Open-Label Extension Study of Study 19
To explore the long-term efficacy and safety of edaravone, investigators conducted an open-label extension of Study 19. A total
of 137 patients who were randomized to receive 60 mg edaravone or placebo for 6 cycles in the double-blind period were offered the opportunity to proceed to the 24-week extension period.

Methods
For inclusion in the extension, patients were required to meet the following criteria: ALSFRS-R score of 2 points or higher, a forced vital capacity of 80% or greater, definite or probable ALS, and duration of disease of less than 2 years.51 Additionally, patients were Grade 1 or 2 ALS severity according to the Japan ALS Severity Classification, aged 20 to 75 years, and had to show a change in ALSFRS-R scoring during the 12-week pre-observation period of –1 to –4 points.51

A minimization method of dynamic randomization was used to balance prognostic factors as patients were randomized to receive either 60 mg edaravone or placebo. In cycle 1, the drug was administered for 14 consecutive days, followed by a 2-week drug-free period. In cycle 2, the drug was administered for 10 days within a 14-day period followed by the same 2-week drug-free period.51

Results
Efficacy. Change in ALSFRS-R score from baseline in the double-blind period was –4.1 ± 3.4 in the edaravone-edaravone (E-E) group and –6.9 ± 5.1 in placebo-edaravone (P-E) group, as compared to –8.0 ± 5.6 in the E-E group and –10.9 ± 6.9 in the P-E group over the 48-week period.51

Safety. Among patients who experienced at least one AE during the extension study, 26.2% in the E-E group and 39.7% in the P-E group were reported as serious. The most common AEs were constipation, dysphagia, and contusion. Of note, no sudden deterioration in ALSFRS-R score was reported in the E-E group, and the investigators detected no safety concerns related to edaravone.51