Individuals using assistive technology may not be able to fully access the information contained in this file. For assistance, please call 800-835-4709 or 240-402-8010, extension 1. CBER Consumer Affairs Branch or send an e-mail to: ocod@fda.hhs.gov and include 508 Accommodation and the title of the document in the subject line of your e-mail.

**FDA Briefing Document**

**BLA 125782**

**Drug name: debamestrocel (MSC-NTF, NurOwn)**

**Applicant: Brainstorm Cell Therapeutics**

**Cellular, Tissue, and Gene Therapy Advisory Committee Meeting**

**09/27/2023**

**Office of Therapeutic Products**
**Center for Biologics Evaluation and Research, FDA**

**DISCLAIMER STATEMENT**

**The attached package contains background information prepared by the Food and Drug Administration (FDA) for the panel members of the Advisory Committee. The FDA background package often contains assessments and/or conclusions and recommendations written by individual FDA reviewers. Such conclusions and recommendations do not necessarily represent the final position of the individual reviewers, nor do they necessarily represent the final position of the Review Division(s) or Office(s). We have brought these issues to this Advisory Committee to gain the Committee's insights and opinions, and the background package may not include all issues relevant to the final regulatory recommendation; instead it is intended to focus on issues identified by the Agency for discussion by the Advisory Committee. The FDA will not issue a final determination on the issues at hand until input from the Advisory Committee process has been considered and all reviews have been finalized. The final determination may also be affected by issues not discussed at the Advisory Committee meeting.**

# Table of Contents

# Table of Tables



# Table of Figures

Glossary

| | |
|---|---|
| ALS | amyotrophic lateral sclerosis |
| ALSFRS-R | ALS Functional Rating Scale-Revised |
| BDNF | brain-derived nerve growth factor |
| BLA | Biologics License Application |
| CAFS | Combined Assessment of Function and Survival |
| CFR | Code of Federal Regulation |
| CI | confidence interval |
| CSF | cerebrospinal fluid |
| DMEM | Dulbecco's Modified Eagle's Medium |
| EAP | expanded access protocol |
| FDA | Food and Drug Administration |
| HGF | hepatocyte growth factor |
| KM | Kaplan-Meier |
| LAP | latency-associated peptide |
| LIF | leukemia inhibitory factor |
| LS | least-squares |
| MCP-1 | monocyte chemotactic factor-1 |
| mITT | modified intent-to-treat |
| MSC-NTF | mesenchymal stromal cells secreting neurotrophic factors |
| NfL | neurofilament light chain |
| NTF | neurotrophic factor |
| RTF | Refuse to File |
| SAE | serious adverse event |
| SD | standard deviation |
| SE | standard error |
| SOD1 | superoxide dismutase 1 |
| SVC | slow vital capacity |
| TEAE | treatment-emergent adverse event |
| VEGF | vascular endothelial growth factor |

# 1. Executive Summary and Draft Points for Consideration by the Advisory Committee

## 1.1 Purpose of the Advisory Committee Meeting

The Cellular, Tissue, and Gene Therapies Advisory Committee will meet on September 27, 2023, to discuss the Biologics License Application (BLA) submitted by Brainstorm Cell Therapeutics (the Applicant) for debamestrocel (mesenchymal stromal cells secreting neurotrophic factors [MSC-NTF], NurOwn) for treatment of patients with amyotrophic lateral sclerosis (ALS).

ALS is a rare, relentless, and ultimately fatal neurodegenerative disorder. ALS is characterized by progressive weakness and atrophy of skeletal muscle. Approximately 90% of ALS cases are sporadic, with the cause unknown; the remainder are genetic. Most patients die within 3 to 5 years after symptom onset, due to respiratory muscle weakness resulting in respiratory failure. Around 10% of patients with ALS survive for 10 or more years. MSC-NTF is an autologous cell-based therapy intended to secrete neurotrophic factors (NTFs), proteins that are produced by a variety of cells in the body and are important for the survival and function of neurons. The Applicant hypothesizes that MSC-NTF could slow disease progression in patients with ALS by secreting NTFs. The treatment requires initial bone marrow aspiration from the patient to obtain the cells for production of the product, followed by repeated intrathecal administration.

On initial receipt of the BLA, FDA determined that the submission was scientifically incomplete to demonstrate substantial evidence of effectiveness, and that the manufacturing information was grossly deficient to ensure adequate product quality. Examples of critical information not provided in the BLA submission include missing or inadequate control of materials, validation of methods missing or incomplete, lack of data demonstrating manufacturing consistency, control strategy for prefilled syringe not provided, inadequate manufacturing and testing facility information, and facilities not ready for inspection.

FDA therefore refused to file the submission and detailed these deficiencies in a Refuse to File (RTF) letter to the Applicant. The Applicant elected to request that the BLA to be filed over protest, and subsequently provided further retrospective analyses and biomarker results.

FDA recognizes the urgent unmet need for additional effective treatments for ALS. At the same time, the critical statutory requirements for approval of drugs and biologics include substantial evidence of effectiveness and evidence of safety, and demonstration of adequate product quality. [1] Although the agency has potential issues with the quality and safety of this product, at this advisory committee meeting FDA is seeking to obtain input from the committee as to whether the available data can be considered to constitute substantial evidence of effectiveness to support regulatory approval of MSC-NTF for treatment of ALS.

## 1.2 Overview of Clinical Development Program

The MSC-NTF clinical development program consists of four clinical studies: two single-arm, open-label, early-phase studies (MSC-NTF-001-IL and MSC-NTF-002-IL); one Phase 2 randomized, double-blind, placebo-controlled study (BCT-001-US), in which subjects received a one-time administration of a single intrathecal injection together with 24 intramuscular injections, of either MSC-NTF or placebo; and one

[1] The Food, Drug and Cosmetic Act, section 505(d) (21 U.S.C. § 355(d))

Phase 3 randomized, double-blind, placebo-controlled study (BCT-002-US), in which subjects received a total of three intrathecal injections (one every 8 weeks) of either MSC-NTF or placebo.

The Phase 3 study (BCT-002-US) is the only controlled study to evaluate administration of MSC-NTF using both the intended route and dose interval.

(1) Study BCT-002-US failed to demonstrate efficacy for the primary efficacy endpoint and all key secondary efficacy endpoints.

   (a) No statistically significant difference was observed for the primary efficacy endpoint, the percent of "responders" in the MSC-NTF group compared to the placebo group: 32.6% (31/95) met that responder criterion in the MSC-NTF group and 27.7% (26/94) in the placebo group; odds ratio: 1.33, 95% confidence interval [CI]: 0.63, 2.80; p-value: 0.45.

   A responder was defined as a subject for whom the slope of the post-treatment ALS Functional Rating Scale–Revised (ALSFRS-R) regression line was ≥1.25 points/month larger than the slope of the pre-treatment regression line.[2]

   (b) For the Combined Assessment of Function and Survival (CAFS) scores at Week 28, the least squares mean was 96.5 for subjects in the MSC-NTF group, versus 93.5 for those in the placebo group (95% CI: -11.4, 17.4; nominal p-value: 0.68[3]).

   (c) For the change in ALSFRS-R total score from baseline to Week 28,[4] the least squares mean change was -5.5 for the MSC-NTF group, versus -5.9 for the placebo group, (95% CI: -1.47, 2.20); nominal p-value: 0.69.[3]

   (d) The same percentage (14%) of subjects in both groups had ≥100% improvement in slope of the post-treatment ALSFRS-R regression line at Week 28, compared to the slope of the pre-treatment regression line.

   (e) For slow vital capacity (SVC), the least squares mean was -12.9 (standard error [SE] 1.8) for the MSC-NTF group, versus -11.6 (SE 1.8) for the placebo group. The least squares mean difference was -1.39, with (95% CI: -6.15, 3.38) and nominal p-value 0.56.[3]

(2) Survival in the Phase 3 study was worse at study completion for subjects who received MSC-NTF. A total of 13 deaths occurred during the post-treatment follow up (28 weeks ± 5 days) with 10 deaths (10/95) in the MSC-NTF group and 3 deaths (3/94) in the placebo group. The Kaplan-Meier (KM) estimate of survival at Week 28 (± 5 days) was 88.3% (95% CI: 79.3, 93.6) for the MSC-NTF group and 94.4% (95% CI: 81.2, 98.4) for the placebo group, with a nominal p-value of 0.04 from unadjusted log rank test.

   This outcome suggests the lack of efficacy of MSC-NTF on survival of patients with ALS.

[2] Refer to Section 2.3 ALS Functional Rating Scale-Revised for details about ALSFRS-R, including assessment with slope.

[3] "Nominal" p-values are calculated without multiplicity protection, and consequently lack interpretability. They are provided here only for completeness.

[4] Missing data was assumed to be missing at random. Missing data was handled implicitly through use of the mixed model for repeated measures (MMRM). Missing data due to deaths were not treated differently than missing data due to discontinuation.

(3) The Applicant performed three different retrospective analyses on an unblinded, post-hoc subgroup from the Phase 3 study, excluding in each certain subjects based on the assertion of a "floor effect" in the ALSFRS-R, according to different criteria. A floor effect refers to insensitivity of an outcome measure to differences at the lower end of an assessment scale. In this case, the Applicant claims that a floor effect results in plateauing of ALSFRS-R total scores over time, during which further deterioration of function cannot be measured. However, no floor effect was demonstrated in the analyses. In addition, floor effect would not be expected in the assessment of survival or biomarkers.

Of note, when assessed by change in ALSFRS-R total score from baseline to Week 28, the MSC-NTF subjects ostensibly affected by a "floor effect" in fact experienced a numerically larger decline in function over time than did the corresponding placebo subjects. This result indicates continued deterioration of function and suggests lack of treatment benefit for MSC-NTF subjects.

(4) In the Phase 3 study, the Applicant collected cerebrospinal fluid (CSF) samples at baseline, Weeks 2, 4, 8, 12, 16, and 20 post-Treatment 1, and examined levels of multiple biomarkers. The Applicant then conducted numerous exploratory analyses, including multiple post hoc analyses, to evaluate the relationships between the selected biomarkers and clinical efficacy outcomes, to support the claim of effectiveness. Of note, there was a large amount of missing data for all biomarkers at Week 20 (~50%), the last time point for biomarker sample collection and the focused time point for biomarker analyses.

To be of utility, changes in biomarkers should correlate (positively or negatively) with disease progression, and a potentially effective treatment should demonstrate a persuasive association between the selected biomarkers and clinical efficacy outcomes. For example, neurofilament light chain (NfL) is released into the CSF by damaged or degenerating axons, and is typically present in higher levels in patients with ALSs as their condition worsens; a potentially effective treatment should demonstrate a clear reduction of NfL in CSF, and be associated with less decline in ALSFRS-R.

(a) For NfL, subjects who received MSC-NTF had about 8.9% reduction in CSF NfL level at Week 20 (when the last CSF samples were obtained) when compared to the placebo group. Of note, subjects having higher reduction in CSF NfL level appeared to experience greater loss of function (indicated by more decline in ALSFRS-R total score from baseline to Week 28), the opposite of what would be expected. These findings could be due to 50% of missing NfL data at Week 20 and relatively overall small changes in NfL in MSC-NTF group. Either way in the setting of negative Phase 3 study results, the findings related to NfL do not appear to provide direct evidence on treatment effect through changes in NfL.

(b) In addition to NfL, the Applicant selected galectin-1, latency-associated peptide (LAP), monocyte chemotactic factor-1 (MCP-1), and vascular endothelial growth factor (VEGF)-A. However, no evident association was observed between their percent change from baseline at Week 20 and change in ALSFRS-R total score from baseline to study completion at Week 28.

Exploratory subgroup analyses, even when post hoc and thereby lacking scientific rigor, may still yield insights to guide subsequent clinical studies (e.g., generating hypotheses for future testing).

As in this case, however, such subgroup analyses following overall nonsignificant tests in the overall population must always be interpreted with caution (see Appendix X for a primer on the issues with

subgroup analyses in general). They pose a high risk of generating false positive findings, for several reasons: they do not include statistical controls for testing multiple hypotheses; breaking randomization may lead to imbalance in measured and unmeasured baseline prognostic factors; and unblinding enables selection of analyses yielding more favorable results. Consequently, such exploratory analyses provide little confidence on which to base regulatory decisions, and though they may serve to generate hypotheses for testing in future trials, cannot compensate for negative results in a well-controlled study.

This briefing document presents FDA review team's major concerns that the available data do not meet the statutory standard of substantial evidence of effectiveness to support use of MSC-NTF for the treatment of ALS.

### 1.3 Product Quality

In addition to the above clinical and statistical concerns, the review team has substantial concerns about product manufacturing. Those issues have yet to be resolved. To receive a biologics license, an Applicant must establish product quality, and manufacturing and testing facilities must be licensed. Evaluation of the Applicant's manufacturing data and product control strategy are still ongoing. Thus, for the purposes of the Advisory Committee discussion, this briefing document focuses on clinical, statistical, and biomarker analyses.

### 1.4 Draft Points for Consideration

The Advisory Committee is asked to consider the following issues:

(1) Whether substantial evidence of effectiveness meeting the approval standard had been demonstrated by available evidence.

(2) If substantial evidence of effectiveness meeting the approval standard has not been demonstrated by available evidence, please discuss potential designs for a trial to demonstrate substantial evidence of effectiveness for MSC-NTF for the treatment of ALS.

## 2. Introduction and Background

### 2.1 Background of the Condition/Standard of Clinical Care

ALS is a rare, fatal neurodegenerative disease characterized by relentlessly progressive weakness and atrophy of skeletal muscles. These symptoms result from a combination of loss of motor neurons in the brain (upper motor neurons) and in the brainstem and spinal cord (lower motor neurons).[5]

Symptoms typically begin focally, with weakness of a limb, or, in about one-third of patients, with difficulty speaking or swallowing. Progressive neuromuscular respiratory failure is the most common cause of death in ALS. The median survival from the time of symptom onset is 3 to 5 years; about 10% of patients with ALS live 10 or more years.[6]

Frontotemporal dementia may be associated with ALS in 15-50% of cases.

[5] Elman, L, L McCluskey, and C Quinn, 2023, Clinical features of amyotrophic lateral sclerosis and other forms of motor neuron disease, Food and Drug Administration, accessed, 2023, www.uptodate.com/contents/clinical-features-of-amyotrophic-lateral-sclerosis-and-other-forms-of-motor-neuron-disease#.

[6] Chio, A, G Logroscino, O Hardiman, R Swingler, D Mitchell, E Beghi, BG Traynor, and C Eurals, 2009, Prognostic factors in ALS: A critical review, Amyotroph Lateral Scler, 10(5-6):310-323.

The prevalence of ALS is estimated at 7.7 per 100,000 persons in the United States.[7] Around 90% of ALS cases are sporadic, with the remainder familial. For the latter, multiple genes have been identified, including *C9orf72*, *SOD1* (superoxide dismutase 1), *TARDBP/TDP-43* (transactive response DNA-binding protein 43), and *FUS* (fused in sarcoma/translated in liposarcoma). The precise pathogenesis of ALS remains largely elusive, although multiple pathways appear to be involved, including RNA processing, protein aggregation, oxidative stress, mitochondrial dysfunction, autophagy, excitotoxicity, inflammation, and axonal transport. Environmental factors may also play a role.

Three drugs for ALS (sporadic or genetic) have received FDA approval via the traditional approval pathway: riluzole, edaravone, and sodium phenylbutyrate/taurursodiol (Relyvrio). Effectiveness of all three drugs was demonstrated in randomized, double-blind, placebo-controlled studies. For riluzole, two studies showed a moderate increase in survival in patients treated with riluzole compared to placebo. Patients treated with edaravone experienced less decline in function, as measured by the ALS Functional Rating Scale-Revised (ALSFRS-R).[8] The same effect was observed in the clinical study assessing sodium phenylbutyrate/taurursodiol; in addition, exploratory analysis of long-term survival suggested longer median overall survival.

Tofersen recently received FDA approval via the Accelerated Approval pathway for treatment of ALS associated with mutation of the *SOD1* gene. Accelerated Approval was based on the surrogate endpoint of reduction in blood levels of NfL, a biomarker associated with axonal injury and neurodegeneration. Evidence included reduction at Week 28 in plasma NfL (67% difference in geometric mean ratios for tofersen compared to placebo, nominal $p<0.0001$), and CSF SOD1 (34% difference in geometric mean rations for tofersen compared to placebo, nominal $p<0.0001$). In addition, consistently favorable trends for tofersen on clinical endpoints were observed in the open-label extension study. Although the single pivotal study failed to demonstrate a statistically significant treatment difference on the prespecified primary clinical endpoint, in additional post hoc, exploratory analyses, nominally significant improvements with separation over time were noted in both the ALSFRS-R and survival for patients originally randomized to tofersen compared to patients originally randomized to placebo.[9]

Beyond the FDA-approved therapies for the disease itself, several pharmacologic treatments are available to help manage disabling symptoms associated with ALS. Muscle cramps may be ameliorated by mexiletine. An FDA-approved combination of dextromethorphan and quinidine is effective in reducing uncontrolled emotional outbursts (pseudobulbar affect). First-line treatment for excessive salivation (sialorrhea) involves anticholinergic medications such as atropine, hyoscyamine, amitriptyline, glycopyrrolate, and transdermal scopolamine; if these medications are inadequate or poorly tolerated, injection of botulinum toxin into the salivary glands is another option.

Patients with difficulty swallowing may obtain hydration and nutrition via placement of a percutaneous gastrostomy tube.

For patients experiencing respiratory insufficiency, noninvasive positive pressure ventilation is recommended. The implanted Diaphragm Pacing System received FDA approval for patients with ALS with more severe respiratory symptoms. This device is indicated for treatment of chronic hypoventilation in patients who retain sufficient respiratory function (forced vital capacity of greater than 45% of the predicted value); both the right and left parts of the diaphragm muscle must be

[7] Mehta, P, J Raymond, R Punjani, M Han, T Larson, W Kaye, LM Nelson, B Topol, O Muravov, C Genson, and DK Horton, 2023, Prevalence of amyotrophic lateral sclerosis in the United States using established and novel methodologies, 2017, Amyotroph Lateral Scler Frontotemporal Degener, 24(1-2):108-116.

[8] Refer to Section 2.3 ALS Functional Rating Scale-Revised for details about ALSFRS-R, including assessment with slope.

[9] FDA, 2023, Integrated Review, NDA, tofersen, https://www.accessdata.fda.gov/drugsatfda_docs/nda/2023/215887Orig1s000IntegratedR.pdf.

stimulable, as demonstrated by voluntary contraction or by nerve conduction testing of the phrenic nerves. When respiratory compromise becomes severe, survival can be maintained by tracheostomy and permanent invasive mechanical ventilation. In the United States, 5-10% of patients with ALS choose this option.

## 2.2 Pertinent Regulatory History

### 2.2.1 Refuse to File

Upon receipt of a marketing application for a drug or biologic, FDA conducts an initial review to determine whether the application contains the information required according to the Food, Drug, and Cosmetic Act (FD&C Act); the Public Health Service Act, section 351; and FDA regulations (e.g., 21 CFR 314.50 for New Drug Applications, 21 CFR 601.2 for Biologics Licensing Applications). If substantive deficiencies are present, or the review team has concerns that cannot readily be rectified, such as absence of evidence of effectiveness satisfying those statutes and regulations, FDA can refuse to file (RTF) the application.

Applications and supplements accepted for filing should be sufficiently complete to permit a meaningful review. RTF is an important regulatory tool to help FDA avoid unnecessary review of incomplete applications. Incomplete submissions can lead to multiple-cycle reviews and inefficient use of CBER resources and may also delay the review of more complete submissions from other applicants. CBER also believes an RTF action can allow an applicant to begin repair of critical deficiencies in the submission far sooner than if the deficiencies were identified much later in a complete review action and may lead to more rapid approval of safe and effective products.[10]

### 2.2.2 Key Regulatory History

To expedite development of MSC-NTF for treatment of ALS, FDA has had numerous interactions with the Applicant throughout product development (Please see Appendix I for detailed information).

The Applicant submitted the BLA on September 9, 2022. FDA conducted a filing review and determined that a substantive review could not be performed, because the BLA submission was scientifically incomplete and grossly deficient. Critical clinical and manufacturing deficiencies were identified. For clinical, the completed randomized, placebo-controlled clinical studies failed to show efficacy in their prospectively specified efficacy endpoints to demonstrate required substantial evidence of effectiveness. For manufacturing, the required Chemistry, Manufacturing, and Controls information covering several critical categories was not included in the application, and the level of information included was insufficient to perform a full assessment of product quality. Consequently, FDA issued a refuse-to-file letter to the Applicant on November 8, 2022.

The Applicant and FDA met on January 11, 2023. The Applicant presented additional post hoc analyses purporting to show efficacy of MSC-NTF in subjects who were not affected by Applicant-defined "floor effects" of the ALS Functional Rating Scale–Revised. FDA discussed the Applicant's two possible options moving forward:

(1) Submit a new BLA containing all necessary information, including information identified in the refuse-to-file letter of November 8, 2022: FDA noted the requirements to generate evidence of effectiveness, and evidence of safety, from adequate and well-controlled clinical investigations; submit a complete Chemistry, Manufacturing and Controls module with all necessary validation

[10] FDA CBER SOPP 8404: Refusal to File Procedures, https://www.fda.gov/media/111632/download.

studies completed; and have commercial manufacturing facilities that are ready for inspection; or

(2) Request that the BLA be filed over protest.

The Applicant chose the latter, and the BLA was filed over protest on February 7, 2023.

## 2.3 ALS Functional Rating Scale-Revised

The most widely-used outcome measure for assessing functional deficits in patients with ALS in clinical trials is the ALS Functional Rating Scale–Revised (ALSFRS-R). The ALSFRS-R is an ordinal scale consisting of 12 items, covering four domains: bulbar (brainstem), respiratory, gross motor ability, and fine motor ability. Each item is scored on a 5-point rating scale, from 0 (complete dependence) to 4 (normal function). The total score is the sum of the 12 item scores, and ranges from 0 to 48; higher scores indicate better function.

### 2.3.1 Considerations Regarding Use of ALSFRS-R

While the ALSFRS-R is a validated and reliable tool, several important issues should be considered regarding its use in clinical studies, such as those in BLA 125782:

(1) To assess efficacy of MSC-NTF, the Applicant modeled decline on the ALSFRS-R as a linear function and compared the slope of the resulting lines before and after subjects received treatment. Modeling progression as a linear function, however, is not accurate: the disease course for individual patients is heterogeneous, and the pace of progression may vary among patients, as well as within time periods for an individual patient. Short intervals of plateau or even improvement on the ALSFRS-R are not uncommon (Figure 1).[11,12] For example, one report observed that over a 6- to 18-month period, some patients with ALS did not show decline in ALSFRS-R total scores; over a 6-month interval, some patients even demonstrated improved scores (Table 1).

**Table 1. Patients With ALS in the PRO-ACT Database Who Did Not Show Decline, or Showed Improvement, in ALSFRS-R Total Score Over Various Time Intervals**

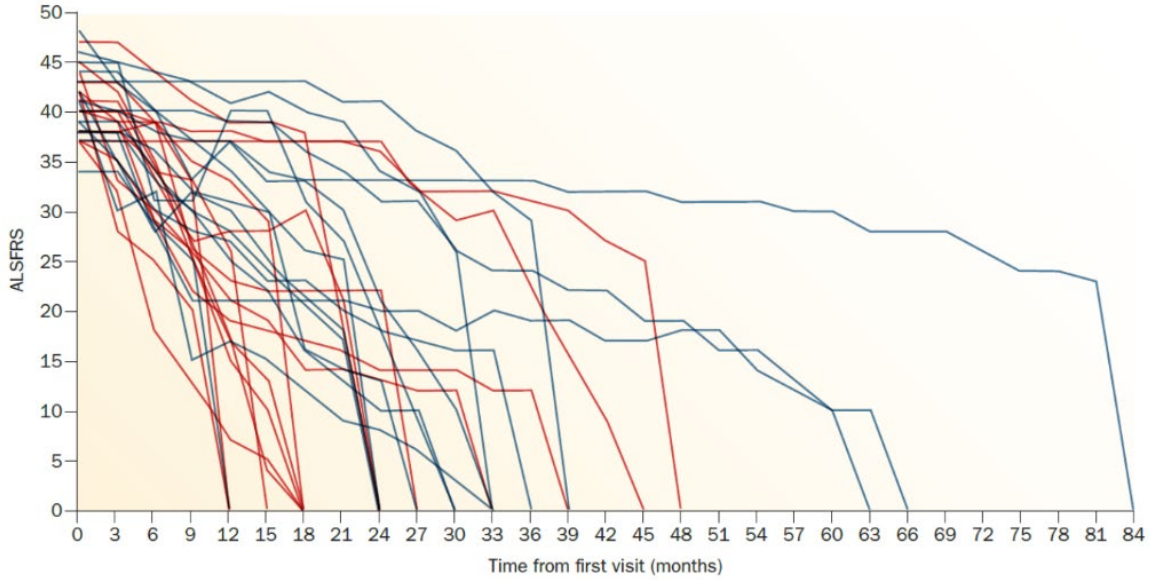
| Time Interval (months) | % No Decline (n) | % Improved (n) |
|---|---|---|
| 6 | 25% (3,132) | 14% (1,343) |
| 12 | 16% (2,105) | — |
| 18 | 7% (1,218) | — |

Source: FDA Clinical Team generated this table based on data from Bedlack, RS, T Vaughan, P Wicks, J Heywood, E Sinani, R Selsov, EA Macklin, D Schoenfeld, M Cudkowicz, and A Sherman, 2016, How common are ALS plateaus and reversals?, Neurology, 86(9):808-812.
Abbreviations: ALS, amyotrophic lateral sclerosis; ALSFRS-R, ALS Functional Rating Scale-Revised; PRO-ACT, Pooled Resource Open-Access ALS Clinical Trials.

[11] Swinnen, B and W Robberecht, 2014, The phenotypic variability of amyotrophic lateral sclerosis, Nat Rev Neurol, 10(11):661-670.

[12] Bedlack, RS, T Vaughan, P Wicks, J Heywood, E Sinani, R Selsov, EA Macklin, D Schoenfeld, M Cudkowicz, and A Sherman, 2016, How common are ALS plateaus and reversals? Neurology, 86(9):808-812.

**Figure 1. Heterogeneity of Decline on the ALSFRS for Individual Patients Over Time**




Source: Swinnen, B and W Robberecht, 2014, The phenotypic variability of amyotrophic lateral sclerosis, Nat Rev Neurol, 10(11):661-670.
Abbreviation: ALSFRS-R, ALS Functional Rating Scale-Revised.

(2) The ALSFRS-R is an ordinal scale. Consequently, the differences between successive points on the ALSFRS-R are not intrinsically equal: a 1-point decrease from 4 to 3 on an item does not necessarily indicate the same extent of functional loss as a drop from 1 to 0.

(3) Scores for different domains within the ALSFRS-R may decline differently. For example, gross and fine motor function may be lost more rapidly than bulbar or respiratory function. Such differences are not readily identifiable when considering changes in the ALSFRS-R Total Score.

(4) Bounded scales such as the ALSFRS-R may be subject to "ceiling" or "floor" effects, where the scale is not able to detect further improvement or deterioration of function. The ALSFRS-R is known to be less sensitive at lower scores.

(5) Dropout or death of subjects in clinical trials results in absence of additional ALSFRS-R data, which may confound interpretation of results by incorrectly implying lack of further disease progression. For example, in a trial in which a proposed treatment actually hastens mortality, the change in ALSFRS-R score for the treatment group may appear better than that for the control group, which experienced fewer deaths but lower ALSFRS-R scores. Joint-rank analyses (e.g., the CAFS[13] and the ALS/SURV[14]) have been developed to account for these factors in studies using the ALSFRS-R.

### 2.3.2 Applicant's Analyses: Linear Regression Modeling of Change in ALSFRS-R Over Time, and Use of Slope

To assess the effect of treatment, the Applicant utilized two different criteria to compare the percent of study subjects in the MSC-NTF and placebo groups who were considered "responders." The following is

[13] Berry, JD, R Miller, DH Moore, ME Cudkowicz, LH van den Berg, DA Kerr, Y Dong, EW Ingersoll, and D Archibald, 2013, The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials, Amyotroph Lateral Scler Frontotemporal Degener, 14(3):162-168.
[14] Goutman, SA, MB Brown, M Cudkowicz, N Atassi, and EL Feldman, 2019, ALS/SURV: a modification of the CAFS statistic, Amyotroph Lateral Scler Frontotemporal Degener, 20(7-8):576-583.

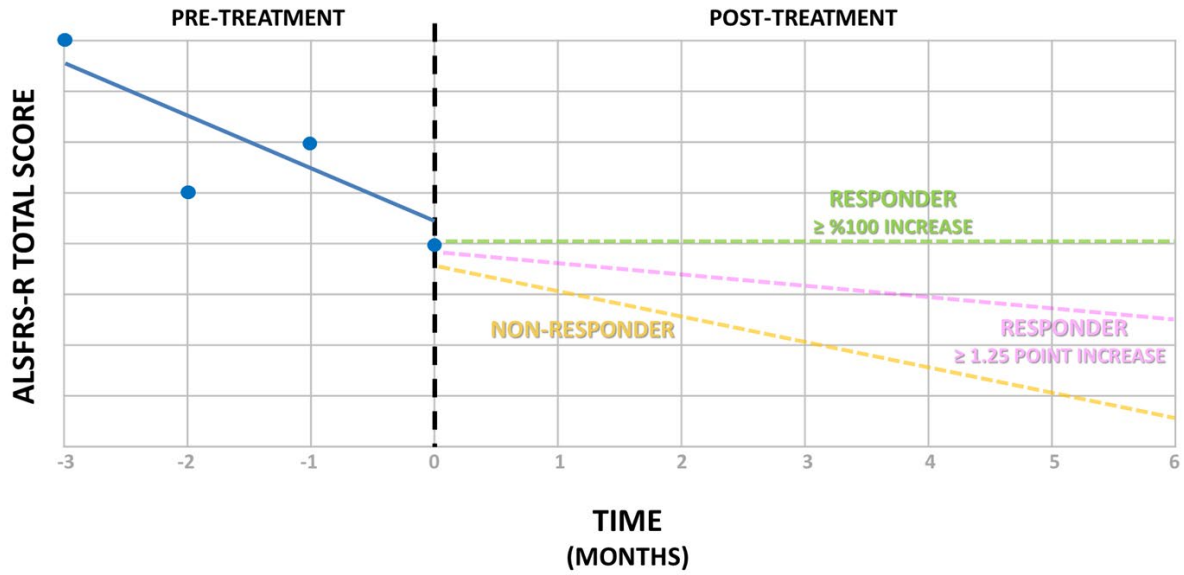
a brief overview of the calculations the Applicant conducted to determine whether a subject qualified as a "responder."

For each subject, the Applicant performed least-squares (LS) linear regression analysis on ALSFRS-R results from the 12 weeks from screening to randomization (i.e., prior to treatment with either MSC-NTF or placebo), and a separate LS linear regression analysis on ALSFRS-R results from the 28 weeks after initiation of treatment. Subjects only qualified for inclusion in the Phase 3 study if the slope of the calculated pre-treatment line was ≤-1 point/month (considered by the Applicant as "rapid progressors").

To determine response to treatment, the Applicant compared the slope of the calculated post-treatment line to the slope of the calculated pre-treatment line:

- For the primary efficacy endpoint of the Phase 3 study, subjects were considered "responders" if the slope of the post-treatment line was ≥1.25 points larger than the slope of the pre-treatment line (e.g., a change from -1 points/month before treatment to 0.25 points/month or above to Week 28 after starting treatment). Subjects not meeting this criterion were considered "non-responders." Subjects who died during the study were also considered "non-responders."
- A secondary efficacy endpoint of the Phase 3 trial considered subjects as "responders" only if the slope of the post-treatment line was ≥100% larger than the slope of the pre-treatment line (e.g., a change from -1 points/month before treatment to 0 points/month or above after treatment). The Applicant interpreted this change as indicating that the subject's disease progression was "halted or improved." Subjects not meeting this criterion were considered "non-responders."

**Figure 2. Hypothetical Illustrations of How Subjects Are Categorized Based on Comparison of Slope of Post-treatment Regression Line to Slope of Pre-Treatment Regression Line**




Source: FDA Clinical Review Team
Note: Yellow line represents a hypothetical "non-responder" subject under both criterion (in this case, the slope of the pre -treatment regression line is steeper than the slope of the post-treatment regression line). Pink line represents a hypothetical "responder" subject under the primary efficacy endpoint criterion (the slope of the post-treatment regression line is ≥1.25 larger than slope of the pre-treatment regression line). The green dashed line represents a hypothetical "responder" subject under the secondary efficacy endpoint criterion (the slope of the post-treatment regression line is ≥100% larger than the slope of the pre-treatment regression line).
Abbreviation: ALSFRS-R, ALS Functional Rating Scale-Revised.

In Figure 2, the solid blue line on the left represents linear regression performed on hypothetical data from a subject's pre-treatment ALSFRS-R total scores prior to the beginning of the treatment period

(initiation of treatment is indicated by the black vertical dashed line). On the right are examples of a hypothetical study subject who would be considered a "non-responder" under both criteria (yellow dashed line); a hypothetical study subject considered a "responder" under the Phase 3 study primary efficacy endpoint criterion (pink dashed line); and a hypothetical study subject considered a "responder" under the secondary efficacy endpoint criterion (green dashed line).

### 2.3.3 FDA Perspective on Use of ALSFRS-R in Clinical Studies

ALSFRS-R Linear Regression Slope

Change in "slope" of ALSFRS-R is of unclear clinical significance, and therefore would not be suitable as the primary efficacy endpoint in a study intended to provide primary evidence of effectiveness in support of a marketing application. As discussed earlier, modeling ALS progression as a linear function based on change in ALSFRS-R over time is inaccurate, because short intervals of plateau or even improvement are not uncommon. Moreover, an acceptable clinical outcome measure should reflect whether a patient feels or functions better or survives longer. A change of 1 point or more in the ALSFRS-R total score compared to a patient's baseline corresponds to a clinically meaningful difference in functional ability; in contrast, the clinical meaningfulness of a change in the slope of the regression line is unknown.

ALSFRS-R and Survival

As noted above, functional endpoints such as ALSFRS-R can be confounded by loss of data because of patient deaths. To address this issue, FDA recommends an analysis method such as a joint-rank test, that combines ALSFRS-R and survival into a single overall measure. The CAFS and ALS/SURV are examples of such joint-rank measures.[15]

For trials with few or no deaths, change from baseline in the ALSRFS-R total score may be an acceptable efficacy endpoint.

# 3. Investigational Product

## 3.1 Drug Product Description and Dose

Debamestrocel is composed of mesenchymal stromal cells, also referred to as mesenchymal stem cells. It is an autologous product generated from a single bone marrow aspirate. The manufacturing process is intended to increase levels of NTFs normally produced by MSC to generate a cellular product referred to by the Applicant as MSC-NTF cells. Although the product is a cellular product, this is not intended as a cell replacement strategy. The product is intended to work through secretion of factors by MSC-NTF cells into the surrounding CSF after transplantation. The product is formulated at a final cell concentration of $25 \times 10^6$ cells/ml to $31.25 \times 10^6$ cells/ml in 4 ml of Dulbecco's Modified Eagle's Medium (DMEM). The product is provided as a combination product in a prefilled capped syringe. Each 5 ml syringe is shipped directly to the health care site or administration center in a temperature-controlled shipping system to maintain a temperature between 2 to 8°C. The proposed expiry (shelf life) is 48 hours. Enough cells are intended to be manufactured from a single bone marrow collection to produce all three treatment doses. The product is administered into the CSF using a 20-gauge spinal needle into the L3/L4 or L4/L5 lumbar intervertebral space.

[15] FDA, 2019, Amyotrophic Lateral Sclerosis: Developing Drugs for Treatment: Guidance for Industry, www.fda.gov/regulatory-information/search-fda-guidance-documents/amyotrophic-lateral-sclerosis-developing-drugs-treatment-guidance-industry.

The intended dose is $125 \times 10^6$ cells. If less than $125 \times 10^6$ cells are produced, the total available cell dose shall be administered, provided at least $100 \times 10^6$ cells are administered. For the Phase 3 BCT-002-US study, subjects received either $100 \times 10^6$ cells or $125 \times 10^6$ cells; there was no intermediate dose level tested. The full dose was used in 76% of subject treatments. However, limited data are available on the efficacy of the minimum $100 \times 10^6$ dose, as only 6% of subjects treated in the Phase 3 BCT-002-US study received all three doses at that level. Furthermore, it should be noted that there may be manufacturing variability associated with the generated doses due to the differences in the amount of cells present from donor source material and the manufacturing process (see Section 3.3 Product Quality and Manufacturing Concerns for information on the manufacturing process). The impact of product dose on clinical outcome is unclear.

## 3.2 Proposed Mechanism of Action (MOA) of MSC-NTF for Treatment of ALS

The mechanism of action (MOA) of MSC-NTF is unknown. The Applicant's proposed MOA in the BLA is unclear and inconsistent across the BLA. The MSC-NTF product manufacturing control strategy for the Phase 3 study and the intended commercial process consists only of neurotrophic factor secretion. NTFs are proteins that play a critical role in the survival, differentiation, maturation, and neurite outgrowth of developing peripheral and CNS neurons. They also support mature neurons and can contribute to neural plasticity. They are expressed in tissues that are innervated by specific neuronal populations, and by neurons and glial cell within the CNS. They play a fundamental role during nervous system development in matching the appropriate level of neurons with their target tissues. Cultured primary neurons will die if NTFs are withdrawn. NTFs were first identified about 50 years ago, and several families of proteins with neurotrophic properties have been identified. Proteins that have been found to support motor neurons through in vitro or in vivo preclinical testing include members of the neurotrophin family (nerve growth factor, brain-derived nerve growth factor [BDNF], neurotrophin 3, neurotrophin 4), ciliary neurotrophic factor, glial cell line-derived neurotrophic factor, and VEGF. They exert their effect by binding to cell surface receptors expressed on specific populations of neurons.

Due to their biological properties on motor neurons, NTFs have been considered a promising potential therapy for treatment of several neurodegenerative diseases, including ALS. Unfortunately, clinical studies involving administration of purified NTFs have been hampered by delivery limitations in vivo, difficulties in providing a sustained concentration, rapid turnover of the proteins, and in some cases serious side effects.

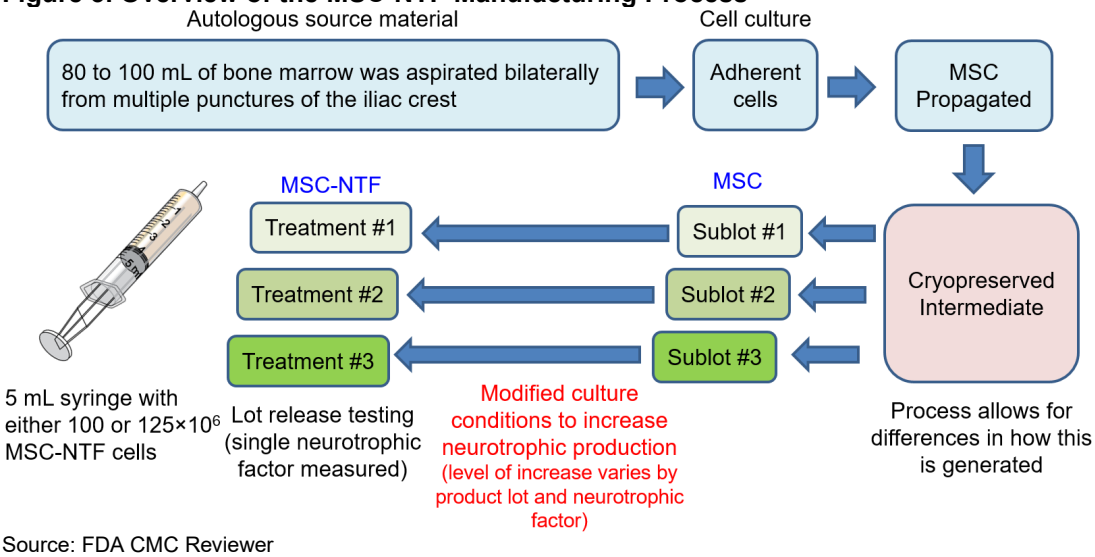
Cell therapies have been proposed as a potential alternate means to provide trophic factors in vivo for neurological and other medical conditions. In theory, cells that continuously secrete NTFs could provide a more sustained source, and the same cells could potentially secrete multiple beneficial molecules. Gene modified cells and viral vector-based gene therapies are another potential means to deliver specific NTFs in vivo. There are currently no licensed cell therapies in the U.S. for the treatment of neurologic disease.

## 3.3 Product Quality and Manufacturing Concerns

The autologous MSC-NTF product is generated from a single bone marrow collection. A total of 80 to 100 mL of bone marrow is aspirated bilaterally from multiple punctures of the iliac crest of the pelvic bone (~5 mL from each puncture) from each subject. The number of bone marrow cells collected varies by collection. The bone marrow cells are shipped to a product manufacturing site. The product was manufactured for the BCT-002 clinical study at two facilities: Dana-Farber Cancer Institute Cell Manipulation Core Facility and the City of Hope - Center for Biomedicine and Genetics. Six clinical sites were used for the Phase 3 study, four of which were in close geographical location to one of the two

manufacturing sites. At the manufacturing site adherent cell cultures are established to begin the manufacturing process (see Figure 3).

**Figure 3. Overview of the MSC-NTF Manufacturing Process**




Source: FDA CMC Reviewer
Abbreviations: MSC, mesenchymal stromal cells.

The adherent cells are grown under conditions to enrich for MSC. The MSC are frozen as a product intermediate to be able to accommodate the timing of production of three individual product lots, each intended to be provided fresh to the subject for treatment at 0, 8, and 16 weeks relative to baseline assessment. Generation of the frozen intermediate can occur by 3 different procedures. Vials of MSC intermediate are thawed and the cells cultured under specific culture conditions intended to elevate the levels of NTFs produced by MSC and to further expand cell numbers. At the end of the incubation period the cells are harvested and formulated in a minimum culture medium and loaded into a 5 ml syringe. A 48 hour expiry (shelf life) is proposed.

FDA has concerns about the consistency of the manufacturing process and potential sources of product variability. It is important for licensure the Applicant demonstrate the manufacturing process is under a state of control. Chemistry, Manufacturing, and Control regulations are intended to assure that all subjects receive a quality product lot, including for safety and potency. Data supporting a product can come from in-process and final product properties, and from clinical data of safety and efficacy. However, for this BLA clinical data supporting safety for all patients is unclear, and efficacy has not been demonstrated. This places greater emphasis on product data. Due to the absence of a large amount of manufacturing information and product data for this file, our assessment has proved challenging.

In addition to concerns about the adequacy of the existing manufacturing control strategy, there are concerns about manufacturing changes – either those that occurred during clinical development under IND, or for the proposed commercial product. In general, manufacturing changes can occur at any point in the product lifecycle, though it is best to introduce significant changes early in product development before clinical studies intended to support safety and efficacy are conducted. Further, FDA discourages incorporation of changes during BLA review. Depending on the extent and type of change, such changes may raise questions about how representative the pre-change product is to the to-be-licensed commercial product. A comparability study may be necessary for certain types of changes. It can be challenging to establish comparability through analytical methods. In cases where analytical methods alone cannot sufficiently establish comparability, additional preclinical or clinical studies may be needed. If the Applicant were to conduct a new clinical study, such as, but not limited to, the now-proposed

Phase 3/4 clinical study, we would recommend all manufacturing changes be incorporated prior to conducting the study so that clinical data would exist using the post-change product.

Concerns about manufacturing variability and product comparability include:

(1) Comparability of product quality across different clinical studies included in the BLA to support safety and efficacy was not demonstrated. Production of product lots for the Phase 1/2 BCT-001-IL study, BCT-001-IL Phase 2a study, BCT-001-US Phase 2b study, Phase 3 BCT-002-US study, and BCT-003-US Expanded Access program involved different manufacturing conditions, facilities, and specifications. Comparability among these clinical study products was not assessed by the Applicant, nor was comparability between the two Phase 3 manufacturing sites.

(2) Control of product variability for the Phase 3 BCT-002-US study was not demonstrated. No information was provided on the source material or in-process step yields for the Phase 3 BCT-002-US study. It is therefore difficult to determine how consistent the upstream processing steps were for the product used in the Phase 3 study.. The manufacturing process allows for the frozen intermediate to be generated using 3 different procedures. It is unclear whether the same level of product quality is achieved using all three variants of the manufacturing process. The Applicant did not identify which process was used to generate each product lot, and it was not indicated which subjects received a product lot made by which process.

(3) Process validation has not been performed. Process validation is critical for demonstrating the consistency of the manufacturing process, for identifying sources of variability, and for establishing process limits at critical steps in the manufacturing process.

(4) The intended dose for the Phase 3 BCT-002-US study was $125 \times 10^6$ cells; however, if less than $125 \times 10^6$ cells were available, the total available cell dose was to be administered, provided at least $100 \times 10^6$ cells were administered. Approximately a quarter of the product doses generated did not achieve $125 \times 10^6$ cells. This included product lots made from the same autologous frozen intermediate where other doses did achieve the full dose. It is not clear why the manufacturing process failed to achieve the full dose for all three patient-specific lots in cases where other $125 \times 10^6$ cell doses were produced. A single collection is intended to provide enough cells that will expand during the manufacturing process to generate all 3 treatment doses.

(5) The level of NTF production is highly variable across product lots. Further, which NTFs were assessed for product release varied by clinical study, and in the Phase 3 BCT-002-US study only a single NTF was measured. Data from the BCT-001-US Phase 2b study where multiple factors were measured on each lot (though not as part of lot release), showed the amount of individual NTFs varied considerably by product lot, and by individual NTF.

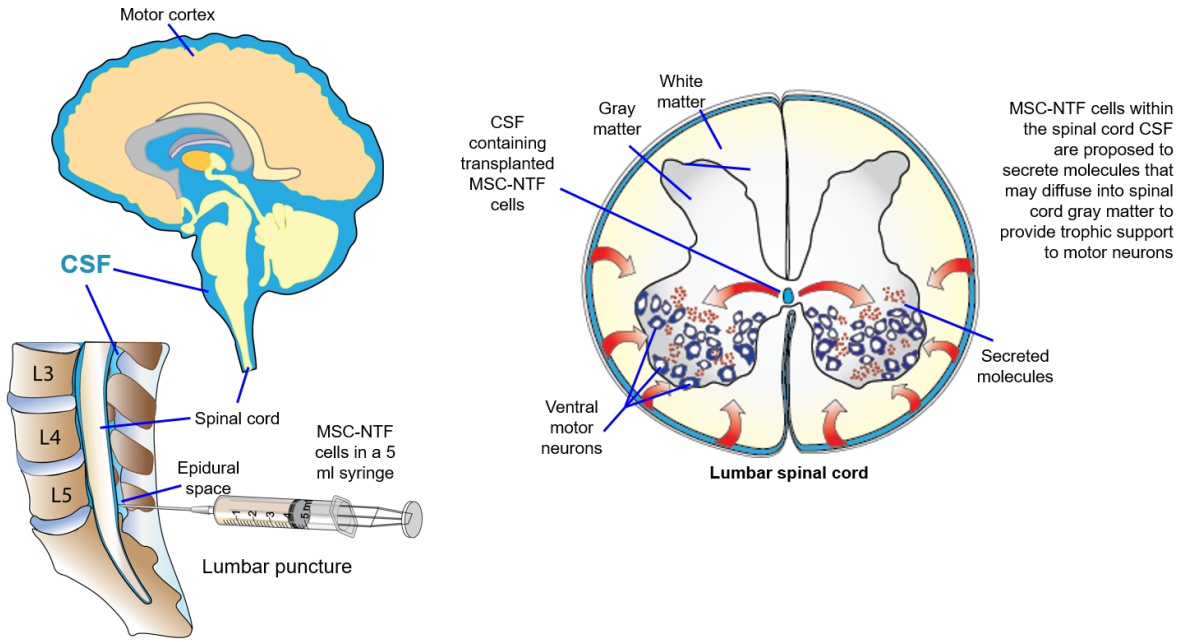
(6) The Applicant states that mechanisms other than NTF secretion may be involved in the function of the product in vivo and may involve secretion of molecules other than NTFs. However, no manufacturing control strategy exists for any other potential mechanism of action.

(7) Additional concerns were identified regarding the intended commercial process. These include a change in a critical manufacturing reagent and in manufacturing facilities without supporting comparability data. There are also concerns about the adequacy of product stability data in prefilled syringes.

## 3.4 Assessment of the Proposed MOA

The product is delivered into the CSF at the L5 lumbar spinal level (see Figure 4). CSF surrounds the spinal cord and is present in the central canal that runs the length of the spinal cord. It also surrounds the brain and fills the four ventricles. CSF is generated by the choroid plexus in the lateral, third and fourth ventricles, and flows throughout the CNS. The production rate in humans is 0.3–0.6 ml/min, turning over a total CSF volume of 150 ml 4 times per day[16]. MSC-NTF cells transplanted within the spinal cord CSF are proposed to secrete neurotrophic molecules. ALS results in loss of motor neurons at different levels in the spinal cord and in the motor cortex of the brain. Molecules secreted by MSC-NTF cells and released into the CSF may diffuse to other spinal cord or brain regions. It might also diffuse into spinal cord gray matter to ventral motor neurons, thereby providing trophic support.

**Figure 4. Route of Administration and Proposed Mechanism of Action for MSC-NTF for ALS**




Source: FDA CMC Reviewer
Abbreviations: ALS, amyotrophic lateral sclerosis; CSF, cerebrospinal fluid.

It is unclear how much diffusion of proteins released by MSC-NTF occurs after transplant. It is also unclear if the cells move to other regions of the spinal cord or brain, either through bulk flow of the CSF or by migration. It is also unclear if transplanted cells integrate directly into CNS tissue and if so, in what regions. Since ALS affects motor neurons through all levels of the spinal cord and the motor cortex of the brain, migration of cells or diffusion of secreted molecules throughout CSF would be beneficial.

Because the product is intended to function through elevation of NTFs in the CSF, FDA was particularly interested in evaluating CSF biomarker data for the four NTFs included in the biomarker panel: BDNF, leukemia inhibitory factor (LIF), hepatocyte growth factor (HGF), and VEGF. We evaluated whether there was any correlation between product properties and the concentration of these four molecules present in CSF samples. We also compared the amount of trophic factor tested for release with levels observed in vivo. Although the level of a secreted NTF was measured for product release, the rate of secretion, and how long the secreted level is maintained in culture was not provided in the BLA. Neurotrophic

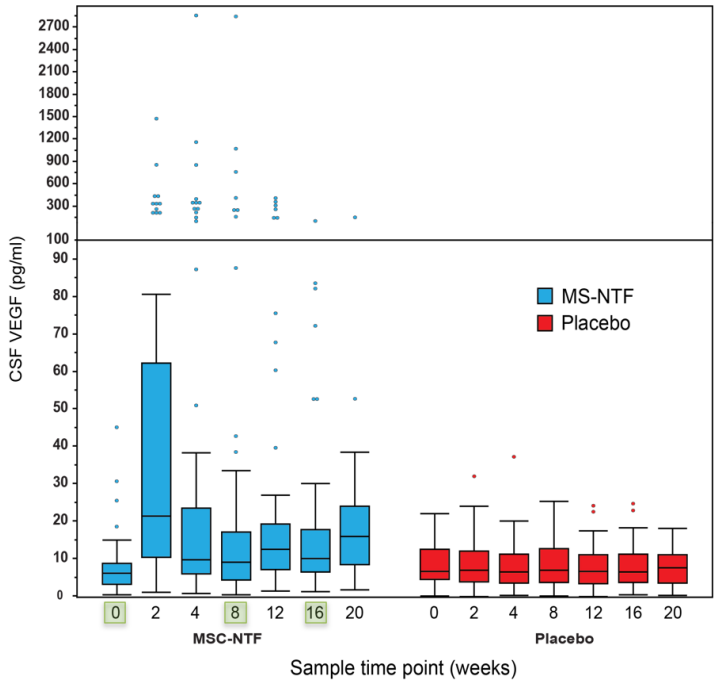
[16] Simon MJ, Iliff JJ. Regulation of cerebrospinal fluid (CSF) flow in neurodegenerative, neurovascular and neuroinflammatory disease. Biochim Biophys Acta. 2016 Mar;1862(3):442-51. doi: 10.1016/j.bbadis.2015.10.014. Epub 2015 Oct 22. PMID: 26499397; PMCID: PMC4755861.

production was tested on product samples taken from MSC-NTF cells cultured under conditions optimized to increase neurotrophic factor production. The production level by MSC-NTF cells under conditions simulating the composition of CSF does not appear to have been tested. We noted that the panel of neurotrophic molecules measured by biomarker analysis was different than that measured for product release for the Phase 3 BCT-002 trial; only a single factor was measured for potency in the Phase 3 study. We found no correlation between product release properties and either ALSFRS-R clinical scores, or the level of CSF NTFs measured as biomarkers. This lack of correlation makes it challenging to assess whether the manufacturing control strategy is adequate for assuring product quality or predicting the impact of a product change or product manufacturing deviation.

In our evaluation of CSF NTF levels, we noted the low levels of neurotrophic molecules found. It is not known how much of any specific NTF would be needed to provide clinical benefit for patients with ALS. We focused our analysis of CSF biomarkers primarily on VEGF CSF data because the observed increase relative to placebo concentrations for each of the four factors was greatest for VEGF. The median VEGF CSF concentrations were in the low pg/ml range. The median levels reported at all time points for the MSC-NTF treatment group are of questionable pharmacological relevance given what is known about purified VEGF biological activity (see Appendix II).

We also noted extensive variation in VEGF levels at different time points in different subjects. A small subset of total samples accounted for much of the variation, with some individual values well above the median (Figure 5). As part of our evaluation of VEGF biomarker data we chose 100 pg/ml as a cutoff to separate those sample and subjects that had CSF values for VEGF well above the median.

**Figure 5. Variation in VEGF Concentrations in CSF Samples Associated With a Subset of Subjects Who Had Elevated Levels at One or More Time Points**



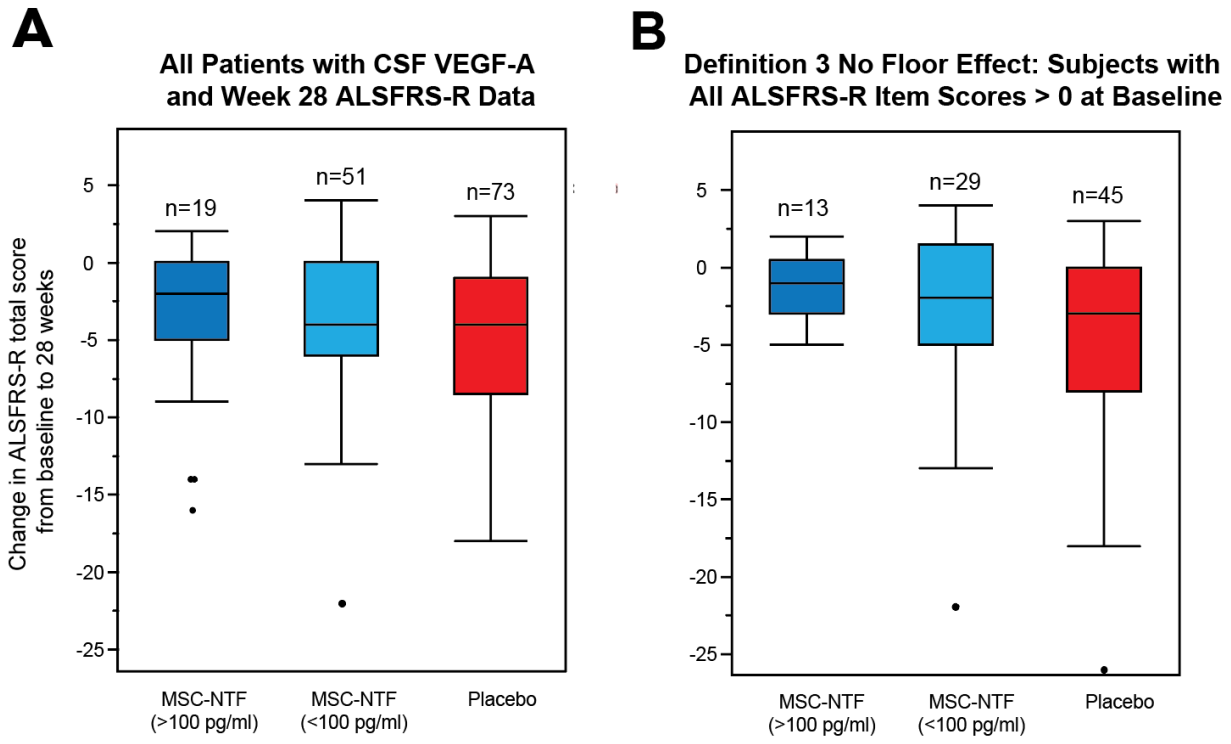

Source: FDA CMC Reviewer
Abbreviations: CSF, cerebrospinal fluid; VEGF, vascular endothelial growth factor.

Most MSC-NTF samples at all time points were <100 pg/ml. Only ~8% of samples across all time points had concentrations >100 pg/ml, 73% of subject samples did not exceed 100 pg/ml at any time point, and no subject had CSF samples that exceeded 100 pg/ml at all 6 time points post initiation of treatment. Of

the 24 subjects where higher levels were seen, only 4 had levels elevated at 3 or more of the 7 time points measured. The peak observed at 2 weeks across all subject samples was not observed in all individual subjects, and others had peak levels at different points in the timeline that did not seem to correspond well with timing of product administration. It was also unclear in the case of a peak at 8 weeks, why a peak would be seen only at 8 weeks and not at earlier time points after the first infusion, or how levels could become elevated at the time of the second administration before secreted molecules could accumulate in the CSF. In terms of the CSF VEGF data supporting the presence of MSC-NTF cells after product administration, the peak observed at 2 weeks could be an indication that the cells do not persist for long periods of time. Subsequent CSF samples were taken at 4-week intervals, and those levels were closer to placebo levels, and with fewer examples of levels >100 pg/ml.

Since the proposed MOA is to elevate levels of trophic molecules in the CSF with the hope of reducing disease progression through sparing of motor neurons, we correlated the clinical outcome of those subjects with VEGF CSF concentrations ≥ 100 pg/ml to see if those subjects with higher concentrations achieved less disease progression. For our analysis we included any subject with elevated CSF NTF at any of the 7 collection time points. We saw no correlation between a change in ALSFRS-R scores from baseline to 28 weeks in the elevated VEGF CSF group (see Figure 6A) compared to patients that had low VEGF CSF values at all time points collected. We applied the same analysis using the Applicant's Definition 3 floor effect criteria. There was less variation, due to the smaller sample size, but no correlation with improved disease progression (Figure 6B). These data suggest either that the VEGF concentrations, even at the highest levels observed at any time point, are not sufficient to impact clinical outcome, or that the duration in vivo is too brief.

**Figure 6. Higher CSF VEGF Levels Do Not Correlate With Better ALSFS-R Results**

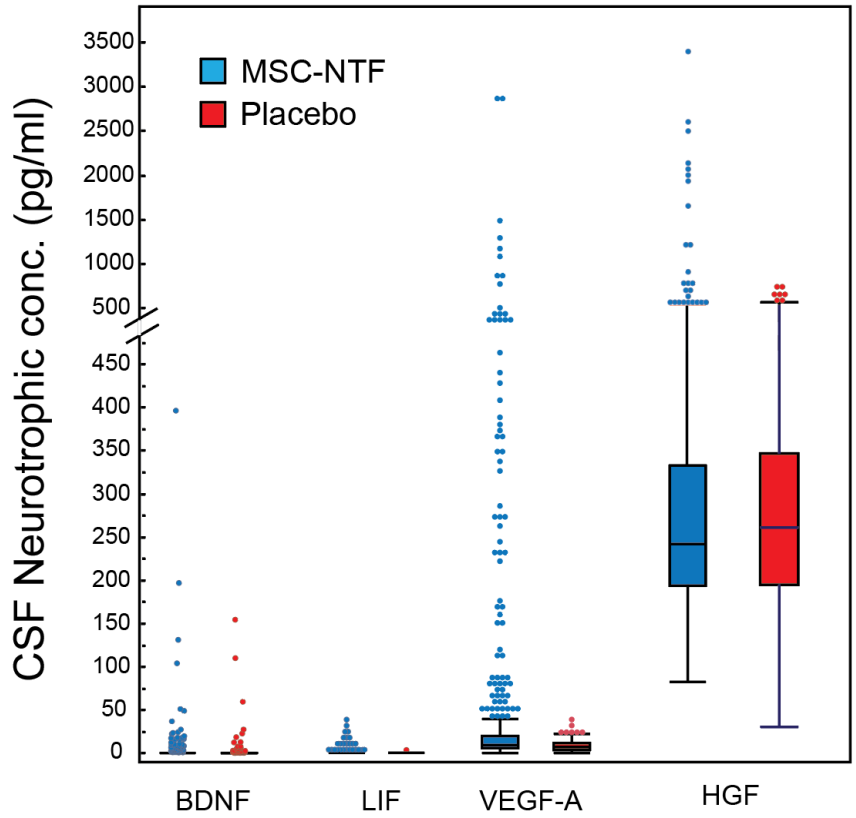



Source: FDA CMC reviewer
Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised; CSF, cerebrospinal fluid; mITT, modified intent-to-treat population; VEGF, vascular endothelial growth factor.

With regards to data on BDNF, LIF, and HGF, we observed the following (see Figure 7):

- Of the four neurotrophic molecules measured in the CSF, VEGF stood out as the most increased compared to placebo, but the increase was mostly due to a fraction of samples that were outliers. Median values across all post-administration time points and patients showed little overall increase: 12.2 pg/ml MSC-NTF (n=407) versus 7.1 pg/ml for placebo (n=416)
- BDNF and LIF concentrations were very low. For example, in the MSC-NTF samples the median value of all samples 2 weeks or greater was 0.12 pg/ml (n=393) and 0.08 pg/ml (n=406) for the placebo. Many of the data points were below the lower limit of quantitation of the BDNF assay.
- HGF was already expressed at significant levels in placebo subjects and only a small number of treated subjects had higher levels

**Figure 7. Comparison of CSF Levels of BDNF, LIF, VEGF, and HGF Among All Collection Time Points**




Source: FDA CMC Reviewer
Abbreviations: BDNF, brain-derived nerve growth factor; CSF, cerebrospinal fluid; HGF, hepatocyte growth factor; LIF, leukemia inhibitory factor; VEGF, vascular endothelial growth factor.

Complicating the interpretation of CSF NTF data:

- About 20% of samples were missing
- Many samples were not collected precisely according to clinical schedule – timing of sample collection relative to product administration is important for NTFs because the persistence of MSC-NTF cells after administration is unclear.
- How far the cells move from the injection site might conflate measured CSF VEGF levels – CSF biomarkers are sampled from the same lumbar location as administered.

- Cell analysis from collected CSF samples was not performed to document how many cells were present, if they were viable, their properties, or how long the cells persisted after transplant.

To summarize the NTF data, the proposed MOA does not appear to be well supported due to:

(1) Lack of important data about cell survival and persistence

(2) No correlation between in vitro levels of a neurotrophic factor and CSF levels, or between elevated VEGF levels in vivo and improved clinical outcome

Given the generally low levels of NTFs observed in the CSF, and the incremental improvement over placebo, it does not appear that the use of MSC to deliver NTFs in vivo has overcome the limitations encountered using purified NTFs for the treatment of neurodegenerative disease.

When the mechanism of action is not known it can be challenging to determine whether appropriate measures of product quality are in place for release of each product lot. The more product knowledge that exists (such as critical quality attributes), and the greater the understanding of how the product may be working in the subject, the more supportive data that is available to justify the approach to product quality. A potential advantage of some cell therapies is that they can work through more than one mechanism. However, this can pose challenges in setting appropriate in-process and final product release specifications. For example, it may be necessary to include more than one measure of product potency for product release.

Changing the proposed MOA or adding additional MOAs for a product can occur as additional product and clinical knowledge is gained. In fact, it is a regulatory requirement that release specifications be reviewed on a regular basis to assess whether they are still relevant measures of product quality. Typically, the change in, or addition to, an existing MOA occurs during early-phase studies. It can be problematic to add new MOAs at the BLA stage, as clinical lots used to support the safety and efficacy of the product will likely not include the additional critical quality attributes associated with the new MOAs.

The Applicant has suggested the product may work through other possible MOAs, such as reduction of neuroinflammation, increase in neuroprotection, or through neurogenesis. Study BCT-002-US was not designed to investigate other possible MOAs beyond measurement of a panel of biomarkers. Further, the panel did not include common inflammatory markers (e.g., IL-1, IL-12, IL-17, IL-21, IFNγ, TNFα) or anti-inflammatory markers (e.g., IL-4, IL-6, IL-10) often assessed when inflammation is implicated as a contributing factor in disease, or when clinically evaluating a reduction in inflammation. It is also unclear if the manufacturing process that is optimized for NTF production has a similar impact on the secretion of other types of molecules. Notably, the percent difference observed in CSF samples in the treatment group at 2 weeks post transplantation for MCP-1 was much smaller than in the case of VEGF, and the subgroup of elevated outliers seen for VEGF was not present in the MCP-1 samples. We found no apparent correlation between CSF levels of VEGF and MCP-1 from the same patient (at any time point), and NTF levels measured for release do not correlate with CSF MCP-1 levels. Therefore, measurement of NTF levels for release does not appear predictive of immunomodulatory properties. Importantly, there is no manufacturing control strategy for any biologic activity other than potency of a single neurotrophic factor. Additional measures of biological activity could be incorporated into a revised manufacturing control strategy, or as part of clinical outcome measures if a new trial is conducted. It is up to the Applicant to determine appropriate measures to assure adequate product quality of every lot.

# 4. Clinical Efficacy

## 4.1 Sources of Clinical Data

Data from four clinical studies are submitted to the BLA. Table 2 summarizes these studies.

The two early-phase studies (MSC-NTF-001-IL and MSC-NTF-002-IL) were single-arm, open-label studies, intended to evaluate the safety and tolerability of MSC-NTF administered intrathecally and/or intramuscularly. The remaining two studies (the Phase 2 study BCT-001-US, and the Phase 3 study BCT-002-US) were randomized, double-blind, and placebo-controlled.

An intermediate-sized expanded access protocol (EAP), BCT-003-US, is ongoing. At the time of the BLA submission, BCT-003-US enrolled 10 subjects, all of whom had participated in the Phase 3 study, to allow continued access of MSC-NTF. The BLA contained limited information regarding BCT-003-US and did not include the study report or datasets. BCT-003-US consequently is discussed only to a limited extent in this Briefing Document.

In addition, 19 subjects have received treatment with MSC-NTF in Israel through a compassionate-use program. The BLA did not include a report of these 19 subjects.

Since ALS is a heterogeneous disorder, and functional outcome assessments such as the ALSFRS-R are effort-dependent, efficacy data from open-label studies are difficult to interpret. Therefore, this Briefing Document focuses on the efficacy data from studies BCT-001-US and BCT002-US.

**Table 2. Clinical Studies Enrolling Patients With ALS**

| Study Identifier | Study Design | Study Objectives | Dosing Regimen | Number of Subjects | Key Eligibility Criteria | Study Duration |
|---|---|---|---|---|---|---|
| MSC-NTF-001-IL | • First-in-human, open-label, single-arm<br>• Two subject groups: "early stage" ALS and "progressive stage" ALS<br>• Single center (Israel) | Evaluate safety and tolerability of a single intrathecal administration of MSC-NTF, together with a single administration of MSC-NTF via multiple intramuscular injections | • Single intrathecal administration of ~60 × $10^6$ MSC-NTF in injection volume of 2 ml<br>• Single intramuscular administration of MSC-NTF injected into 24 separate sites along the right biceps and triceps muscles (~1 × $10^6$ cells per site in 200 µl, for a total of ~24 × $10^6$ cells) | • Planned: 24<br>• Treated: 12 (6 subjects in each group) | • Adult patients with sporadic ALS (revised El Escorial definite or probable)<br>• "Early-stage" ALS defined as ALSFRS-R total score ≥30; disease duration <2 years; sufficient muscle bulk<br>• "Progressive stage" ALS defined as ALSFRS-R total score between 15 to 30; disease duration <2 years; FVC ≥50% of predicted | 3 months pre-treatment, 6 months post-treatment |
| MSC-NTF-002-IL | • Open-label, dose-escalation (3 dose levels)<br>• Subjects with "early-stage" ALS<br>• Single center (Israel) | Evaluate safety, tolerability, and preliminary efficacy of co-administration of intrathecal and intramuscular injections of escalating doses of MSC-NTF | • Single intrathecal administration of MSC-NTF at high, medium, or low dose (70 × $10^6$ cells, 105 × $10^6$ cells, or 140 × $10^6$ cells respectively)<br>• Single intramuscular administration of MSC NTF injected into 24 separate sites along the right biceps and triceps muscles | • Planned: 12<br>• Treated: 14 (4 subjects in low-dose cohort, 6 subjects in medium-dose cohort, and 4 subjects in high-dose cohort) | • Adult patients with sporadic ALS (revised El Escorial definite or probable)<br>• "Early-stage" ALS, defined as ALSFRS-R total score ≥30; disease duration <2 years; sufficient muscle | 3 months pre-treatment, 6 months post-treatment |

| Study Identifier | Study Design | Study Objectives | Dosing Regimen | Number of Subjects | Key Eligibility Criteria | Study Duration |
|---|---|---|---|---|---|---|
| | | | (total of either 24 × $10^6$ cells, 36 × $10^6$ cells, or 48 × $10^6$ cells for high, medium, or low dose, respectively) | | bulk; and FVC ≥ 50% of predicted | |
| BCT-001-US | • Phase 2, randomized (3 MSC-NTF: 1 placebo), double-blind, placebo-controlled<br>• Three centers (United States) | • Primary: Safety<br>• Secondary: Compare difference in pre-treatment slope versus post-treatment slope of ALSFRS-R and of SVC | • Single intrathecal administration of MSC-NTF (~100-125 × $10^6$ cells) or placebo<br>• Single intramuscular administration of MSC-NTF (total of 48 × $10^6$ cells) or placebo, injected into 24 separate sites along the right biceps and triceps muscles | • Planned: 48<br>• Randomized: 48<br>• Treated: 48 | • Adult patients with ALS (revised El Escorial definite, probable, laboratory-supported probable, or possible)<br>• Disease onset >12 months to ≤24 months<br>• ALSFRS-R total score ≥30<br>• Must have limb weakness<br>• SVC ≥65% predicted<br>• Stable dose of riluzole or no history of riluzole treatment | 12-16 weeks pretreatment, 24 weeks post-treatment |
| BCT-002-US | • Phase 3, randomized (1:1), double-blind, placebo-controlled<br>• Multicenter (United States) | • Evaluate efficacy of MSC-NTF compared to placebo (measured by proportion of subjects with difference in pre-treatment slope versus post-treatment slope of ALSFRS-R of ≥1.25) at 28 weeks | • Three intrathecal administrations, 8 weeks apart, of MSC-NTF (100-125 × $10^6$ cells) or placebo in 4 ml per injection | • Planned: 200<br>• Randomized: 196<br>• Treated: 189 | • Adult patients with ALS (revised El Escorial definite, probable, laboratory-supported probable, or possible)<br>• Disease onset, including limb | 18-20 weeks pretreatment, 28 weeks post-treatment |

| Study Identifier | Study Design | Study Objectives | Dosing Regimen | Number of Subjects | Key Eligibility Criteria | Study Duration |
|---|---|---|---|---|---|---|
| | | following first treatment<br>• Evaluate safety of MSC-NTF compared to placebo<br>• Evaluate biomarkers in CSF and serum to assess relationship to treatment | | | weakness, <24 months<br>• ALSFRS-R total score ≥25<br>• Decline in ALSFRS-R total score of ≥3 points in the 12 weeks preceding randomization<br>• SVC ≥65% predicted<br>• Stable dose of riluzole or no history of riluzole treatment | |

Source: Applicant Module 2.5 Clinical Overview
Note: All four studies were completed at the time of BLA submission.
Note: Revised El Escorial diagnostic criteria categorizes patients with ALS into 4 levels of diagnostic certainty (clinically definite, probable, laboratory-supported probable, and possible)
Abbreviations: ALS, amyotrophic lateral sclerosis; ALSFRS-R, ALS Functional Rating Scale-Revised; CSF, cerebrospinal fluid; FVC, forced vital capacity (the volume of air that can be forcibly exhaled from the lungs after taking the deepest breath possible); SVC, slow vital capacity (reported as a percent of normal for gender, height, and age).

## 4.2 Phase 3 Study: Study BCT-002-US

### 4.2.1 Study Overview

The Phase 3 study (BCT-002-US) was a randomized, double-blind, placebo-controlled study in adult subjects with ALS.

Key eligibility criteria included diagnosis of possible, laboratory-supported probable, probable, or definite ALS, based on the revised El Escorial criteria; symptom onset within 24 months; ALSFRS-R score ≥25 at screening; and upright SVC at the screening visit ≥65% of predicted (SVC is reported as a percent of normal for gender, height, and age).

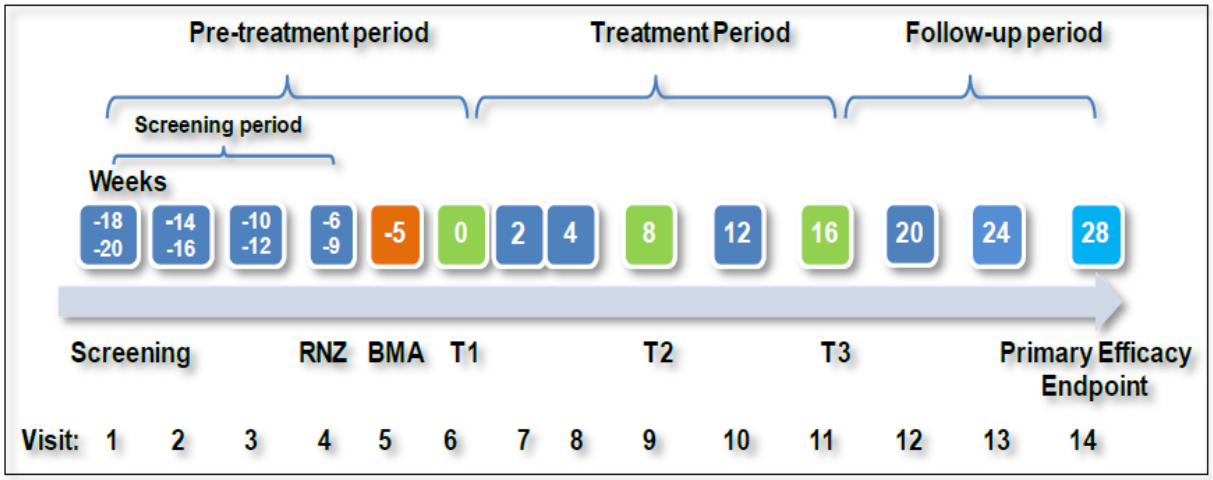
Subjects were scheduled to receive three intrathecal administrations, spaced 8 weeks apart, of either placebo or of MSC-NTF at a dose of 100-125 × $10^6$ cells (Figure 8).

A total of 263 subjects were screened. During the 12-week run-in period, subjects were assessed for decline in the ALSFRS-R; only those subjects who experienced on average at least a 1.0-point monthly decline were included in the trial. Of the subjects screened, 196 were randomized in a 1:1 ratio to receive either MSC-NTF or placebo. For statistical analysis, this group constitutes the intent-to-treat population.

A total of 189 subjects received at least one treatment of either MSC-NTF or placebo. For statistical analysis, these subjects compose the modified intent-to-treat (mITT) population.

The study follow-up duration was 28 weeks (± 5 days) after the first intrathecal administration. The protocol did not include a plan for continued follow-up to comprehensively collect data on survival or functional status after the final study visit.

**Figure 8. Overview of Phase 3 Study (BCT-002-US)**




Source: Applicant's BCT-002-US Clinical Study Report
Abbreviations: BMA, bone marrow aspiration; RNZ, randomization; T, treatment with MSC-NTF or placebo.

Efficacy Endpoints

The primary efficacy endpoint for each subject was the binary determination of whether that subject was a "responder" or a "non-responder."

As described above, a "responder" was defined as a subject for whom the slope of the post-treatment ALSFRS-R regression line was ≥1.25 points/month larger than the slope of the pre-treatment line (e.g., a change from -1 before treatment to 0.25 or above after treatment). Subjects not meeting this criterion

were considered “non-responders.” Subjects who died were also categorized as “non-responders.” The Applicant then compared the percent of responders in the MSC-NTF group to that in the placebo group.

The key secondary efficacy endpoints were:

(1) The binary determination of whether for each subject the slope of the post-treatment regression line was ≥100% larger than the slope of the pre-treatment line (e.g., a change from -1 before treatment to 0 or above after treatment). Subjects meeting this criterion were considered to demonstrate halting or reversal of disease progression and were designated as “responders.” The Applicant compared the percent of responders in the MSC-NTF group to that in the placebo group.

(2) Average change in ALSFRS-R score from baseline to Week 28 for the MSC-NTF group compared to the placebo group.

(3) Average CAFS score from baseline to Week 28 for the MSC-NTF group compared to the placebo group.

(4) Average change in SVC from baseline to Week 28 for the MSC-NTF group compared to the placebo group.

(5) Proportion of subjects with tracheostomy-free survival in the MSC-NTF group compared to the placebo group.

(6) Proportion of surviving subjects at Week 28 in the MSC-NTF group compared to the placebo group.

Statistical Methods

The primary analysis method for the primary efficacy endpoint was logistic regression, adjusting for covariates of treatment group, baseline ALSFRS-R score, duration from onset of symptoms to first treatment, site of onset (limb versus limb and bulbar), riluzole use, and ALSFRS-R slope pre-treatment. Subjects who died were considered non-responders. The same method was used for the analysis of the secondary endpoint, whether a subject’s disease progression was halted or improved as measured by a 100% or greater improvement in post-treatment slope versus pre-treatment slope in ALSFRS-R score at Week 28.

The change in ALSFRS-R score between baseline and Week 28 was analyzed using mixed effects repeated measures with the change in ALSFRS-R score from baseline as the dependent variable and treatment group, visit, baseline ALSFRS-R score, duration from onset of symptoms to first treatment, site of onset (limb versus limb and bulbar), riluzole use, and ALSFRS-R slope pre-treatment as main effect and the interaction between treatment group and visit. The same method was used for the analysis of the secondary endpoint of change from baseline in SVC at Week 28.

The CAFS score between baseline and Week 28 was analyzed using an analysis of covariance model with treatment as a fixed effect and adjusted for covariates baseline ALSFRS-R score, duration from onset of symptoms to first treatment, site of onset (limb versus limb and bulbar), riluzole use, and ALSFRS-R slope pre-treatment.

The analysis of tracheostomy-free survival was a log-rank test, and a Cox proportional hazards model with treatment, baseline ALSFRS-R score, duration from onset of symptoms to first treatment, site of onset (limb versus limb and bulbar), riluzole use, and ALSFRS-R slope pretreatment as covariates. The same method was used for the analysis of survival.

The mITT population was used for all analyses of primary and key secondary endpoints. The mITT population was defined as all subjects who were randomized, treated, and had at least three ALSFRS-R assessments (i.e., one pre-treatment assessment of ALSFRS-R prior to the baseline assessment, a baseline assessment, and one post-treatment assessment). Baseline assessment was the ALSFRS-R assessment at the first treatment (T1) visit (Week 0, Visit 6) prior to treatment.

A sequential testing strategy was used to control the Type I error rate across the multiple key secondary endpoints, with testing of secondary endpoints in the order listed above (if the primary analysis was statistically significant).

### 4.2.2 Subject Disposition and Baseline Characteristics

Of the 196 subjects randomized, 3% (3/98) of the MSC-NTF group and 4% (4/98) of the placebo group discontinued participation in the study prior to receiving treatment.

Of the remaining 189 subjects (the mITT population), 74.7% (71/95) of the MSC-NTF group and 74.5% (70/94) of the placebo group completed the study (i.e., had ALSFRS-R assessed at the Week 28 visit). The mean age was 48.6 years, and subjects were predominantly white (88.9%) and male (67.2%). The mean baseline ALSFRS-R score was 30.9 (standard deviation [SD]: 6.3), with a range from 16 to 46.

Baseline demographic and disease characteristics are shown in Table 3.

The MSC-NTF group had a slightly lower baseline mean ALSFRS-R total score (30.3 MSC-NTF versus 31.4 placebo), and a slightly higher percentage of subjects with baseline ALSFRS-R total score of <35 (73% MSC-NTF versus 66% placebo). The MSC-NTF group also had a higher percentage of subjects who met revised El Escorial diagnostic criteria for definite ALS (53% MSC-NTF versus 36% placebo). Riluzole use at baseline was higher in the MSC-NTF group (68% MSC-NTF versus 60% placebo).

The percent predicted SVC and months from first symptom to first treatment were similar between the two groups.

Overall, the MSC-NTF and placebo groups were comparable; any differences in outcomes are unlikely to have been due to baseline imbalances in the two populations.

**Table 3. Baseline Demographics and Baseline Disease Characteristics (mITT Population)**

| Characteristic | MSC-NTF (N=95) | Placebo (N=94) | Total (N=189) |
|---|---|---|---|
| Age (year) | | | |
| n (missing) | 95 (0) | 94 (0) | 189 (0) |
| Mean (SD) | 48.1 (9.71) | 49.1 (8.38) | 48.6 (9.07) |
| Median | 51.0 | 51.0 | 51.0 |
| Q1, Q3 | 42.0, 56.0 | 44.0, 55.0 | 43.0, 56.0 |
| Min, max | 21, 60 | 22, 60 | 21, 60 |
| Age Group (year) | | | |
| <55, n (%) | 65 (68.4) | 63 (67.0) | 128 (67.7) |
| ≥55, n (%) | 30 (31.6) | 31 (33.0) | 61 (32.3) |
| Gender | | | |
| Female | 27 (28.4) | 35 (37.2) | 62 (32.8) |
| Male | 68 (71.6) | 59 (62.8) | 127 (67.2) |
| Race | | | |
| Asian, n (%) | 5 (5.3) | 7 (7.4) | 12 (6.3) |
| Black or African American, n (%) | 3 (3.2) | 3 (3.2) | 6 (3.2) |
| Native Hawaiian or Other Pacific Islander, n (%) | 0 | 1 (1.1) | 1 (0.5) |
| White, n (%) | 87 (91.6) | 81 (86.2) | 168 (88.9) |
| Other, n (%) | 0 | 2 (2.1) | 2 (1.1) |

| Characteristic | MSC-NTF (N=95) | Placebo (N=94) | Total (N=189) |
|---|---|---|---|
| Ethnicity | | | |
| Hispanic or Latino, n (%) | 5 (5.3) | 3 (3.2) | 8 (4.2) |
| Not Hispanic or Latino, n (%) | 90 (94.7) | 91 (96.8) | 181 (95.8) |
| Baseline ALSFRS-R Score | | | |
| n (missing) | 95 (0) | 94 (0) | 189 (0) |
| Mean (SD) | 30.3 (6.5) | 31.4 (6.1) | 30.9 (6.3) |
| Median | 31.0 | 32.0 | 32.0 |
| Q1, Q3 | 26.0, 35.0 | 27.0, 36.0 | 26.0, 35.0 |
| Min, max | 16, 46 | 17, 42 | 16, 46 |
| Baseline ALSFRS-R Slope | | | |
| n (missing) | 95 (0) | 94 (0) | 189 (0) |
| Mean (SD) | -1.7 (0.7) | -1.6 (0.8) | -1.6 (0.8) |
| Median | -1.6 | -1.4 | -1.5 |
| Q1, Q3 | -2.2, -1.1 | -2.0, -1.1 | -2.2, -1.1 |
| Min, max | -4, 0 | -5, 0 | -5, 0 |
| Baseline SVC (% Predicted) Score | | | |
| n (missing) | 95 (0) | 94 (0) | 189 (0) |
| Mean (SD) | 76.2 (20.9) | 75.0 (19.8) | 75.6 (20.3) |
| Median | 73.7 | 75.4 | 74.5 |
| Q1, Q3 | 60.4, 92.5 | 60.4, 90.5 | 60.4, 91.1 |
| Min, max | 27, 119 | 27, 114 | 27, 119 |
| Baseline ALSFRS-R Score | | | |
| <35, n (%) | 69 (72.6) | 62 (66.0) | 131 (69.3) |
| ≥35, n (%) | 26 (27.4) | 32 (34.0) | 58 (30.7) |
| El Escorial Criteria for ALS | | | |
| Possible, n (%) | 6 (6.3) | 6 (6.4) | 12 (6.3) |
| Laboratory-supported probable, n (%) | 15 (15.8) | 23 (24.5) | 38 (20.1) |
| Probable, n (%) | 24 (25.3) | 31 (33.0) | 55 (29.1) |
| Definite, n (%) | 50 (52.6) | 34 (36.2) | 84 (44.4) |
| ALS medical history: months since diagnosis | | | |
| n (missing) | 95 (0) | 94 (0) | 189 (0) |
| Mean (SD) | 6.8 (4.35) | 6.1 (4.80) | 6.4 (4.58) |
| Median | 5.5 | 4.4 | 5.2 |
| Q1, Q3 | 3.3, 10.1 | 2.8, 8.9 | 3.0, 9.6 |
| Min, max | 0, 20 | 0, 22 | 0, 22 |
| ALS medical history: months since first symptom to first treatment | | | |
| n (missing) | 95 (0) | 94 (0) | 189 (0) |
| Mean (SD) | 19.6 (5.17) | 19.1 (4.90) | 19.4 (5.03) |
| Median | 18.9 | 19.1 | 18.9 |
| Q1, Q3 | 15.7, 24.2 | 14.7, 23.1 | 15.5, 23.2 |
| Min, max | 10, 31 | 8, 29 | 8, 31 |
| ALS medical history: duration since first symptom to first treatment | | | |
| <1.5 years, n (%) | 39 (41.1) | 43 (45.7) | 82 (43.4) |
| ≥1.5 years, n (%) | 56 (58.9) | 51 (54.3) | 107 (56.6) |
| Use of riluzole at baseline | | | |
| Yes, n (%) | 65 (68.4) | 56 (59.6) | 121 (64.0) |
| No, n (%) | 30 (31.6) | 38 (40.4) | 68 (36.0) |

Source: Applicant's Table 6 and 7 abbreviated BCT-002-US CSR.
Abbreviations: ALS, amyotrophic lateral sclerosis; ALSFRS-R, ALS Functional Rating Scale–Revised; max, maximum; min, minimum; mITT, modified intent-to-treat; SD, standard deviation.

### 4.2.3 Efficacy Issues in Detail

#### *4.2.3.1 Primary and All Key Secondary Efficacy Endpoints Failed to Show Efficacy*

The primary efficacy endpoint and all key secondary endpoints failed to demonstrate efficacy of MSC-NTF compared to placebo (see Appendix III).

From the statistical perspective, when the primary efficacy endpoint in a clinical study fails to show statistical significance, the secondary efficacy endpoints cannot be tested with Type I error control.

In accordance with the Agency's discussions with the Applicant (Face-to-Face Meeting, November 18, 2019), however, FDA reviewed all primary and key secondary endpoint results. The Agency did so for several reasons: (1) although at that meeting the Applicant expressed openness to changing the primary efficacy endpoint from a slope-based analysis, FDA recommended against doing so, in order to avoid compromising the integrity of the Phase 3 study, which the Applicant had already initiated; (2) data for the outcome measures recommended by the Agency, such as CAFS or survival, were collected by the Applicant as secondary efficacy endpoints; and (3) FDA's willingness to exercise regulatory flexibility and desire to better inform subjects and stakeholders.

Primary Efficacy Endpoint

For the Applicant's primary efficacy endpoint, the percent of responders in the MSC-NTF group versus the placebo group did not show a statistically significantly difference: the MSC-NTF group had 32.6% (31/95) responders and the placebo group had 27.7% (26/94). The odds ratio after adjusting for the predefined covariates was 1.33 (95% CI: 0.63, 2.80) with a p-value of 0.45.

Key Secondary Efficacy Endpoints

All key secondary efficacy endpoints failed to show efficacy of MSC-NTF. For example, the least squares (LS) mean CAFS scores at Week 28 did not differ significantly between subjects in the MSC-NTF group and those in the placebo group (3.0: 96.5 versus 93.5; 95% CI: -11.4, 17.4; nominal p-value:[17] 0.68). Similarly, there was minimal difference in LS mean change from baseline to Week 28 in ALSFRS-R total score (0.4: –5.5 versus –5.9; 95% CI: -1.47, 2.20; nominal p-value: 0.69).

Particularly concerning was the larger number of deaths in the MSC-NTF group compared to the placebo group (10 [10.5%] versus 3 [3.2%]; hazard ratio: 3.3; 95% CI: 0.87, 12.66).

(1) ≥100% Improvement in ALSFRS-R Slope

The same percentage (14%) of subjects in both groups had 100% improvement in ALSFRS-R slope at Week 28.

The Applicant used 100% or greater improvement in post-treatment slope versus pre-treatment slope in ALSFRS-R score to define halt or improvement of disease progression. As discussed earlier, a small percentage of patients with ALS may experience brief plateaus or even improvements in the ALSFRS-R total score as part of the natural history of the disease. While the underlying cause of this observation is unclear, it does not necessarily indicate halt or improvement of the progressive disease.

(2) Change From Baseline to Week 28 in ALSFRS-R Total Score

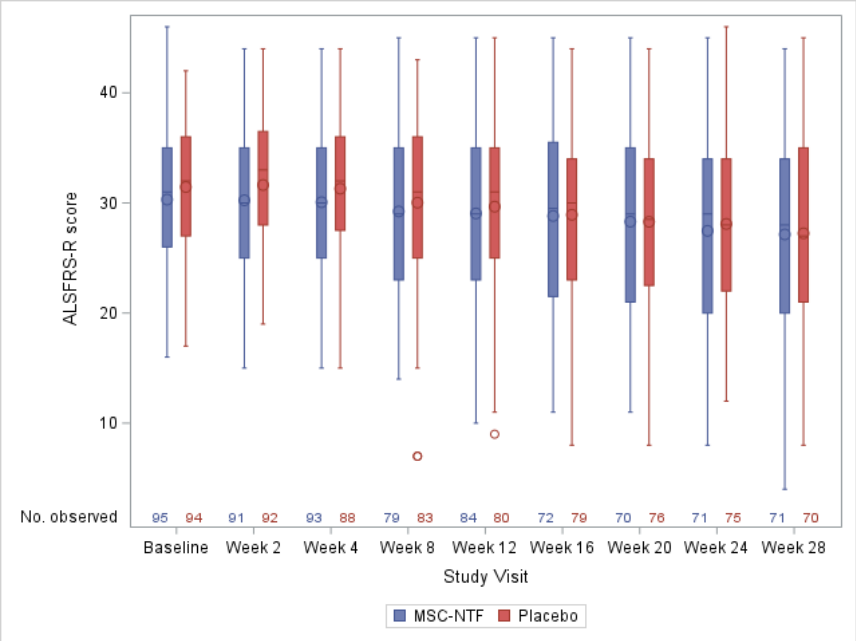
ALSFRS-R total scores were comparable at all time points at baseline, treatment period, and post-treatment follow-up period between the two groups. Both groups demonstrated a consistent decline of ALSFRS-R throughout the study (Figure 9). Figure 9 only shows observed ALSFRS-R score. For subjects

[17] Nominal p-value indicates the lack of multiplicity protection and consequent lack of interpretability.

who died during the study, there were no ALSFRS-R scores available after the death. Considering the higher number of deaths in the MSC-NTF group and these subjects will likely have lower ALSFRS-R scores, the statistics shown in the figure, such as mean and percentiles are likely to be biased and more so for the MSC-NTF group.

The decline in ALSFRS-R scores from baseline to Week 28 was similar between the two groups: LS mean change (SE) from baseline to Week 28 in ALSFRS-R score was -5.52 (0.67) in the MSC-NTF group and -5.88 (0.67) in the placebo group. The maximum LS mean difference was achieved at the Week 12 visit with 1.0 point, and at the other visits, the LS mean difference was within 0.5 point (Figure 10). Changes from baseline in ALSFRS-R is a well-accepted clinical outcome measure in ALS clinical trials. However, such a functional endpoint can be confounded by loss of data because of subject deaths.

**Figure 9. Boxplot of ALSFRS-R Scores Over Study Visit by Treatment Group (mITT Population)**




Source: FDA statistician
Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised; mITT, modified intent-to-treat.

**Figure 10. Change From Baseline in ALSFRS-R Total Score by Week 28 (MMRM Model Over Visit), mITT Population**


Source: FDA statistician
Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised; LS, least squares; MMRM, mixed effects repeated measure; SE, standard error.

(3) Combined Assessment of Function and Survival

The LS mean CAFS (SE) was 96.5 (5.1) for the MSC-NTF group and 93.5 (5.1) for the placebo group. The LS mean difference (SD) was 3.0 (7.3) (95% CI: -11.4, 17.4; nominal p-value: 0.68).

FDA's CAFS analysis was based on the pairwise comparison: "If a participant discontinues early, comparison to each other participant uses time to death if the comparator died before the patient's discontinuation time; otherwise, the comparison is based on the last ALSFRS-R time-point available for both participants."[18] This approach is different from that of the Applicant, where "For subjects whose ALSFRS-R score is missing at a visit, the CAFS rank from the prior visit is used."[19] As a result, FDA's CAFS results are different from those of the Applicant. Despite the difference in the methods to analyzing CAFS, the conclusion of no significant difference between the MSC-NTF and placebo groups remains unchanged.

(4) Slow Vital Capacity

SVC measures the maximum amount of air a subject can exhale in a single breath. SVC is reported as a percent of normal for gender, height, and age. The LS mean SVC (SE) was -12.9 (1.8) for the MSC-NTF group and -11.6 (1.8) for the placebo group. The LS mean difference was -1.39 (95% CI: -6.15, 3.38; nominal p value: 0.56).

[18] Berry, JD, R Miller, DH Moore, ME Cudkowicz, LH van den Berg, DA Kerr, Y Dong, EW Ingersoll, and D Archibald, 2013, The Combined Assessment of Function and Survival (CAFS): a new endpoint for ALS clinical trials, Amyotroph Lateral Scler Frontotemporal Degener, 14(3):162-168.
[19] Goutman, SA, MB Brown, M Cudkowicz, N Atassi, and EL Feldman, 2019, ALS/SURV: a modification of the CAFS statistic, Amyotroph Lateral Scler Frontotemporal Degener, 20(7-8):576-583.

During the COVID-19 pandemic, in-person visits that did not coincide with an administration of treatment were allowed to be converted as remote visits. While remote visits did allow for collection of some clinical data, including the ALSFRS-R, which is validated for remote data collection, SVC was not able to be collected remotely. In addition, most hospitals did not allow the collection of SVC during the COVID-19 pandemic. These COVID-19 pandemic hospital restrictions resulted in 57.9% of the subjects in the MSC-NTF group and 61.7% of the subjects in the placebo group missing Week 28 SVC data. Given the high rate of missing data, SVC is not reported for any subgroup analysis or any post-hoc analysis.

(5) Tracheostomy-free Survival

No subjects underwent tracheostomies during the study. Therefore, tracheostomy-free survival analysis was not done.

#### *4.2.3.2 Survival Was Worse in the MSC-NTF Group*

Survival is the ultimate clinically meaningful outcome measure for a fatal disease like ALS. It is less likely to be affected by variations in assessment. The approval of riluzole was based on its modest survival benefit in over 1,000 subjects with 1-1.5 year follow-up. The challenges for having survival or all-cause mortality as an efficacy endpoint are that survival trials typically require a much larger number of study subjects and longer follow-up time. In recent years, most ALS trials enrolled fewer patients and had shorter post-treatment follow-up time (e.g., 6 months), which makes detecting a survival benefit more challenging. In fact, many ALS trials do not show a difference in survival between treatment and control groups during the short (e.g., 6-month) study follow-up period.

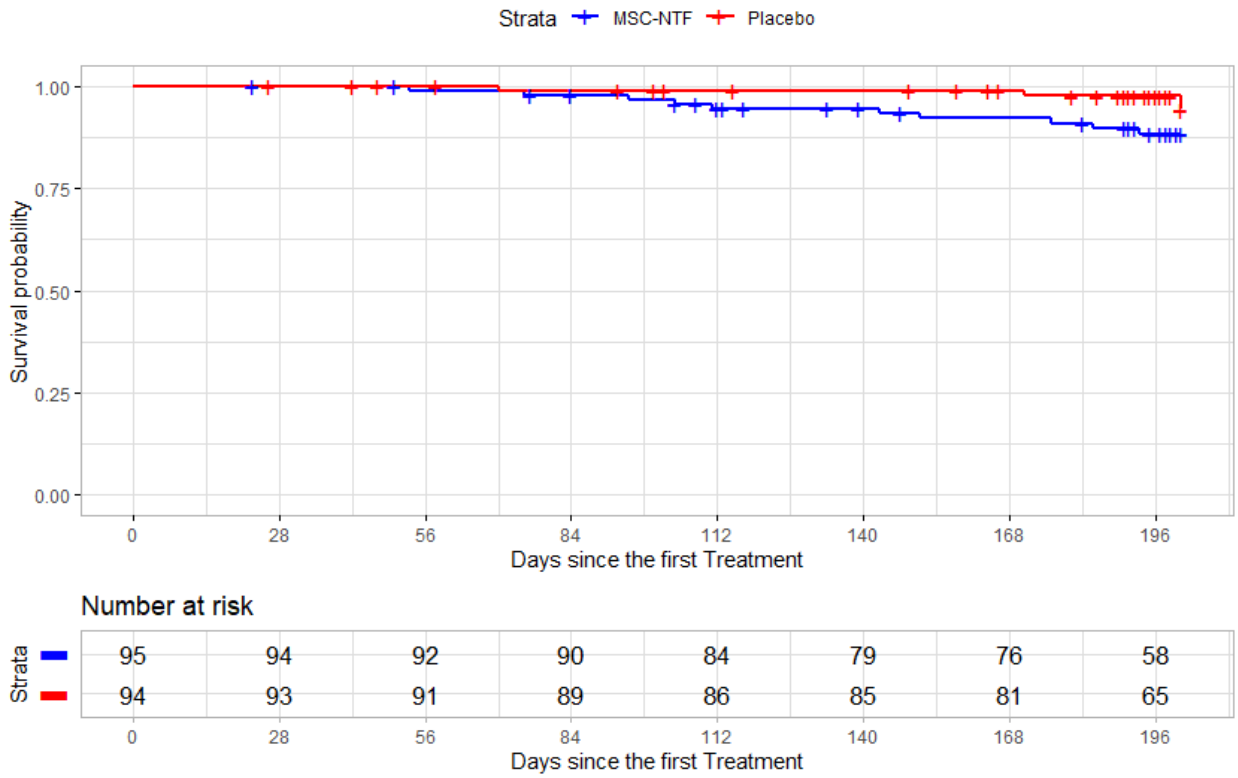
Survival was a prespecified key secondary efficacy endpoint in BCT-002-US. At the study conclusion (Week 28 visit), there were 10 deaths in the MSC-NTF group and two deaths in the placebo group.

It is noted that the Applicant reported two additional deaths in the placebo group that occurred shortly after Week 28.

- The first subject was alive at the final study visit (Week 28 visit). The subject died on Day 201 (28 weeks 5 days) due to an unexpected incident of drowning on Day 196. The cause of death was cardiac arrest and respiratory failure due to drowning as assessed by the investigator. The study protocol allowed a ± 5 days window for each visit. Even though the subject had completed the final assessment at the Week 28 visit, the death occurred within the 5-day window prespecified in the protocol. Therefore, it is reasonable to include the subject in the primary survival analysis.
- The second subject withdrew from the study after the first intrathecal administration of the assigned treatment. The site did an obituary search and found that the subject died after Week 28 (on 29 weeks 2 days). Considering the absence of comprehensive search for the survival status of all subjects who withdrew from the study or completed the study after the protocol allowed final study visit window (i.e., Week 28 + 5 days) and the possibility of bias, this subject is not included in the primary survival analysis.

Notably, the survival endpoint was worse in the MSC-NTF group compared with the placebo group with a hazard ratio of 3.32 (95% CI: 0.87, 12.66). The KM estimate of survival at Day 201 (Week 28 + 5 days) was 88.3% (95% CI: 79.3, 93.6) for the MSC-NTF group and 94.4% (95% CI: 81.2, 98.4) for the placebo group with a nominal p-value of 0.04 from the log-rank test (Figure 11).

**Figure 11. Kaplan-Meier Plot of Overall Survival (mITT Population, Day 201 [Week 28+5 days] Cutoff)**




Source: FDA statistician
Abbreviations: mITT: modified intent-to-treat.

In the BLA submission, the Applicant provided a KM estimate of survival at Week 32. However, as there was no prespecified plan for comprehensive collection of survival data after the final study visit time point (i.e., Week 28 + 5 days), and survival data were available for less than 10% of study subjects after the final study visit, FDA considers it inappropriate to extend the cutoff timepoint to Week 32 when calculating the KM estimate of survival.

The Applicant attributed most deaths to disease progression, which seems reasonable based on available information. However, no autopsy reports were submitted. With the limited information available, it is challenging to assess whether MSC-NTF played an active role in these deaths (e.g., exacerbating disease progression). Nevertheless, the early divergence of the KM curve showing a worsened survival outcome in the MSC-NTF group is of serious concern. Additional well-controlled study(ies) with sufficient number of subjects and duration of follow-up will be critical to further investigate the notably higher all-cause mortality observed in the MSC-NTF group in the Phase 3 trial.

The Applicant reported additional deaths post-study (n=12 in the MSC-NTF group and n=15 in the placebo group). The Applicant became aware of these deaths through a variety of sources (e.g., obtaining consent to use genetic samples, Principal Investigators, media) and captured these data accordingly. However, since there was no systematic long-term follow-up to comprehensively collect survival data of subjects after their completion of or early withdrawal from the study, the actual number of deaths that occurred in each treatment group after study completion is unknown.

#### *4.2.3.3 Subgroup Analyses*

The following subgroup analyses were performed: 1) Subjects with duration since onset of symptoms <1.5 versus ≥1.5 years; 2) Subjects with baseline ALSFRS-R Score <35 versus ≥35; 3) Riluzole use; 4) Sex (male versus female); 5) Race (White versus Black or African American). The Applicant also performed analyses on additional subgroups including age group <55 versus ≥55 years, and site of ALS onset: limb or limb and bulbar.

The Applicant states that these sub-group analyses were "pre-specified." However, it is important to note that these subgroup analyses were not prespecified for hypothesis testing and no prespecified multiplicity adjustment strategy was employed. Subgroup tests following overall nonsignificant tests in the population as a whole such as the Phase 3 study can only be considered exploratory and hypothesis-generating and do not constitute evidence of effectiveness to support marketing approval of MSC-NTF.

Data from exploratory subgroup analyses of primary endpoint suggested possible benefit for MSC-NTF recipients in two subgroups: male subjects with ALS, and subjects with ALS with baseline ALSFRS-R ≥35.

The Applicant considers the finding in the male subgroup spurious. However, the Applicant considers the findings in subjects with ALS with baseline ALSFRS-R ≥35 as evidence of effectiveness. FDA considers that findings from both subgroup analyses could be spurious, as is often the case for such exploratory subgroup analyses. The Applicant may conduct additional well-controlled clinical study(ies) in subjects with high baseline ALSFRS-R to assess the efficacy and safety of MSC-NTF; however, exploratory subgroup analysis data from the completed Phase 3 study cannot serve as evidence of effectiveness.

#### *4.2.3.4 Post-Hoc Floor Effect Analyses Cannot Explain Lack of Efficacy of MSC-NTF*

As described earlier, the ALSFRS-R measures 12 aspects of physical function, ranging from one's ability to swallow and use utensils to climbing stairs and breathing; each function is scored from 4 (normal) to 0 (no ability).

The Applicant argued that "once physical function is lost, and the value of an item reaches 0, further loss cannot be measured even as a patient's condition further deteriorates," and "ALSFRS-R cannot measure further decline once items reach 0, making a treatment effect difficult to measure in participants with lower ratings," and "a floor effect could appear as an improvement or slowing of decline and thereby be misclassified as a clinical response." The Applicant conjectured that the lack of efficacy in the overall population was due to the subgroup impacted by the floor effect.

To support this notion, the Applicant conducted post-hoc analyses based on two types of thresholds to identify subjects not impacted by a floor effect of the ALSFRS-R to define "no floor effect subgroup":

- Definition 1: ALSFRS-R Total Score Threshold with baseline ALSFRS-R >25, which was 77% of the trial participants.
- Definition 2: ALSFRS-R Item Level Threshold with at least two of the six Fine and Gross Motor scale items with baseline values ≥2, which was 84% of the trial participants.

In addition to this analysis, at the Type A meeting with the FDA after refusal to file of the BLA, the Applicant presented a third post-hoc floor effect analysis in which the no floor effect subgroup was defined as ALSFRS-R Item Level had no value 0 at baseline (Definition 3).

We will refer to these subgroups identified by the Applicant collectively as "no floor effect subgroup" and their respective complement "floor effect subgroup."

In the "no floor effect subgroup" identified by different definitions, some of the clinical endpoints showed "statistical significance" per the Applicant (Appendix IV); however, FDA believes these findings

from the exploratory subgroup analysis can only be used for hypothesis generation, not as evidence of effectiveness to support approval, for the following reasons:

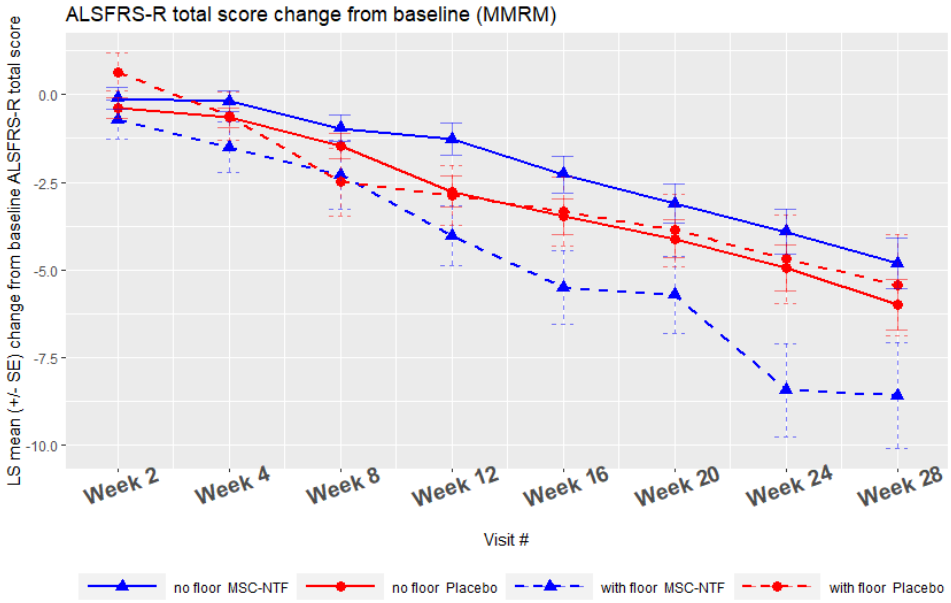
(1) Post-hoc subgroup analyses in general have high risk of finding false positive results due to lack of control for multiple hypothesis testing and potential confounding due to imbalance in the measured/unmeasured baseline prognostic factors brought about by breaking the randomization. What is particularly concerning in this case is that there is no solid definition for the "no floor effect subgroup" (i.e., subgroup of trial subjects not impacted by floor effect). The "no floor effect subgroup" can potentially be defined in many ways, as illustrated by the three distinct subgroups identified by the Applicant, with various sample sizes (145, 159, and 106 subjects respectively). As one could define "no floor effect subgroup" in many ways, some of the "no floor effect subgroup" (like the three selected by the Applicant) may happen to show "positive" findings (i.e., findings that seem to suggest clinical efficacy) among many other subgroups that may show "negative" findings (i.e., findings that seem to suggest harm). These findings could be due to random chance, given the potentially large number of subgroups the Applicant could examine. Therefore, these findings need to be confirmed by additional adequate and well-controlled clinical study(ies) to establish their validity; these findings cannot be used as evidence of effectiveness to meet the statutory standard for this BLA.

(2) MSC-NTF appeared to have a detrimental effect in the floor effect subgroups (Appendix IV). For example, the placebo group had a better CAFS ranking than the MSC-NTF group with a nominal p-value of 0.026 in the floor effect subgroup defined by ALSFRS-R Total Score baseline ≤25 (Definition 1). The floor effect subgroups defined by the other two methods had the same issue. This is not surprising; given that the overall treatment effect was close to zero, when one subgroup happens to show a strong positive treatment effect, the complementary subgroup is highly likely to have a strong negative effect. The "negative" findings in the floor effect subgroup thus may well be false "negative," in the same way that the "positive" findings in the no floor effect subgroup may well be false positives.

(3) FDA did not observe a "floor effect" in the floor effect subgroup defined by any of the three definitions identified by the Applicant. If there were a "floor effect" in the Applicant-identified floor effect subgroup, the ALSFRS-R total score post baseline would have been bounded by a "floor," which would have prevented the score from much further decline. This is in direct contrast with the fact that the MSC-NTF "floor effect subgroup" had a drastically steeper decline in ALSFRS-R total score from baseline compared with the no floor effect subgroup or the placebo floor effect subgroup. At the same time, the magnitude of change between the placebo floor effect subgroup and the placebo no floor effect subgroup were comparable, which further puts into question the validity of the "floor effect" (Figure 12 [using Definition 1] and Figure 14 [using Definition 3]). In addition, the MSC-NTF floor effect subgroup showed substantially worse CAFS ranking than the no floor effect subgroups while the two placebo subgroups were comparable Figure 13.

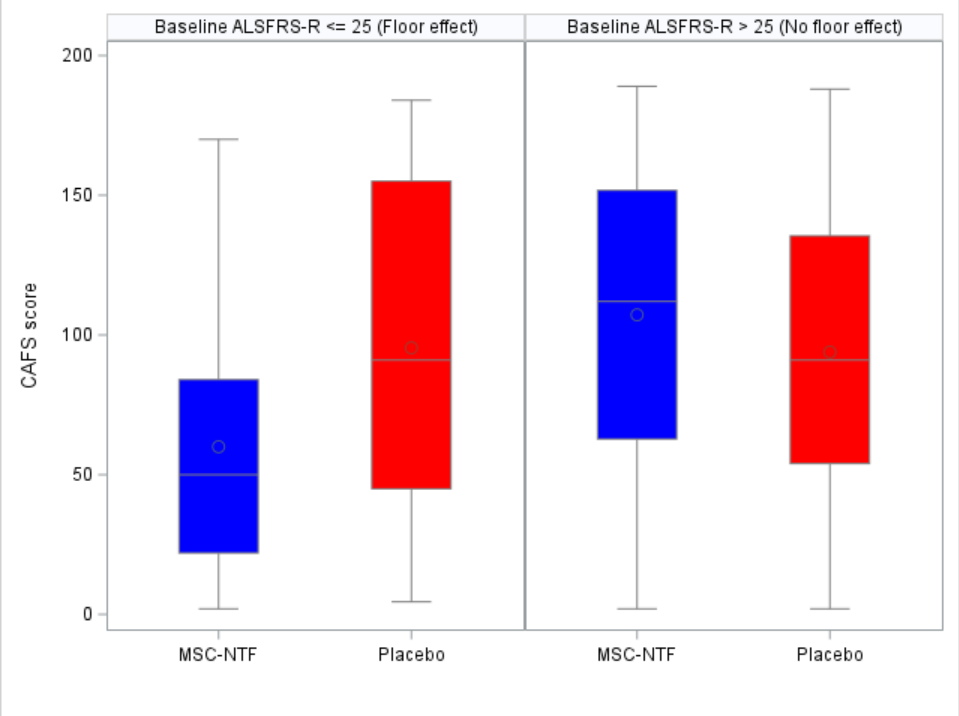
In conclusion, the lack of efficacy of MSC-NTF over placebo cannot be explained by a floor effect.

**Figure 12. Change From Baseline in ALSFRS-R Total Score by Subgroup Based on "Floor Effect" Definition 1 (MMRM Model Over Visit), mITT Population**



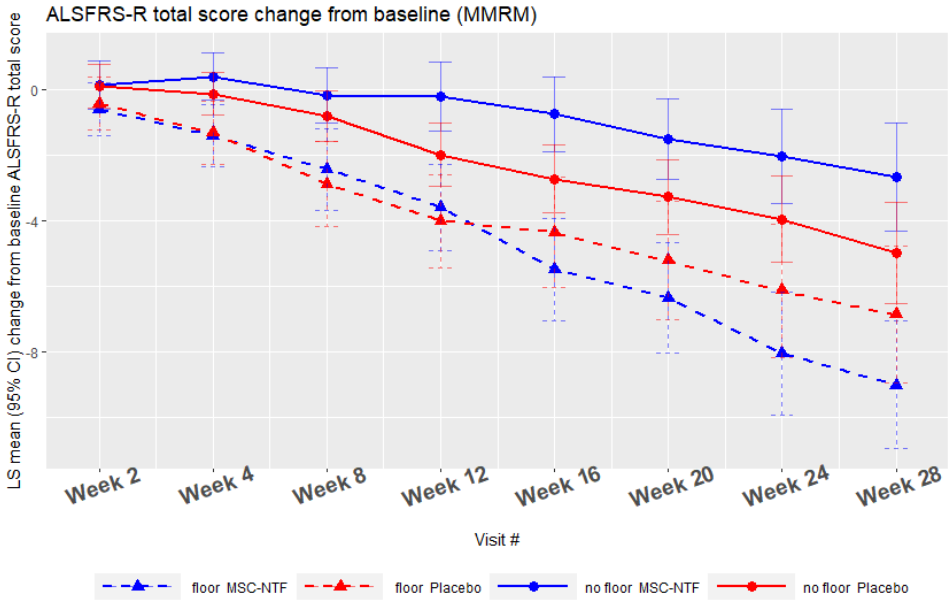

Source: FDA statistician
Abbreviations: ALSFRS-R, ALS Functional Rating Scale–Revised; LS, least squares; mITT, modified intent-to-treat; MMRM, mixed effects repeated measure; SE, standard error.

**Figure 13. Boxplot of CAFS by Subgroup Based on "Floor Effect" Definition 1 and Treatment**




Source: FDA statistician
Abbreviations: ALSFRS-R, ALS Functional Rating Scale-Revised; CAFS, Combined Assessment of Function and Survival.

**Figure 14. Change From Baseline in ALSFRS-R Total Score by Subgroup Based on "Floor Effect" Definition 3 (MMRM Model Over Visit), mITT Population**




Source: FDA statistician
Abbreviations: ALSFRS-R, ALS Functional Rating Scale-Revised; CI, confidence interval; MMRM, mixed effects repeated measure; LS, least squares.

## 4.3 Study BCT-001-US

### 4.3.1 Study Overview

BCT-001-US was a Phase 2, randomized, double-blind, placebo-controlled study to evaluate the safety and preliminary efficacy of MSC-NTF administered on a single occasion via combined intrathecal administration (at 100-125 × $10^6$ cells) and 24 intramuscular injections (at 48 × $10^6$ cells) given into the right biceps and triceps muscles in subjects with ALS. Forty-eight adult subjects with ALS were randomized in a 3:1 ratio to the treatment group (36 subjects) or placebo group (12 subjects). Each subject was followed for approximately 3 months pre-treatment and 6 months post-treatment. Key eligibility criteria included 1) ALS diagnosed as possible, laboratory-supported probable, probable, or definite by revised El Escorial criteria; 2) Disease onset, as defined by first reported occurrence of symptomatic weakness, spasticity, or bulbar symptoms, of more than 12 months and less than or equal to 24 months; 3) ALSFRS-R ≥30 at the screening visit; 4) upright SVC measure ≥65% predicted of normal. The primary endpoint was safety as assessed by the incidence of treatment-emergent adverse events (TEAEs). Secondary endpoints were efficacy evaluated by 1) change in slopes from the pre-treatment period to post-treatment period in ALSFRS-R between treatment and placebo groups through 12 and 24 weeks post-treatment; and 2) change in slopes from the pre-treatment period to post-treatment period in SVC between treatment and placebo groups through 12 and 24 weeks post-treatment. Exploratory endpoints included muscle strength evaluation as performed via hand-held dynamometer, electrical impedance myography, feasibility of blinding, and analysis of CSF.

### 4.3.2 Efficacy Summary

Among the 48 randomized subjects, 43 subjects (89.6%) completed the study and five subjects (MSC-NTF [three subjects, 8.3%] and placebo [two subjects, 16.7%]) discontinued during the treatment period. Demographic characteristics were similar between the two treatment groups. The majority of the treated subjects were male (72.9%). All subjects were white. The mean age of subjects was 51.1 years (range: 26 to 71 years). The use of riluzole was allowed and the most frequently reported prior

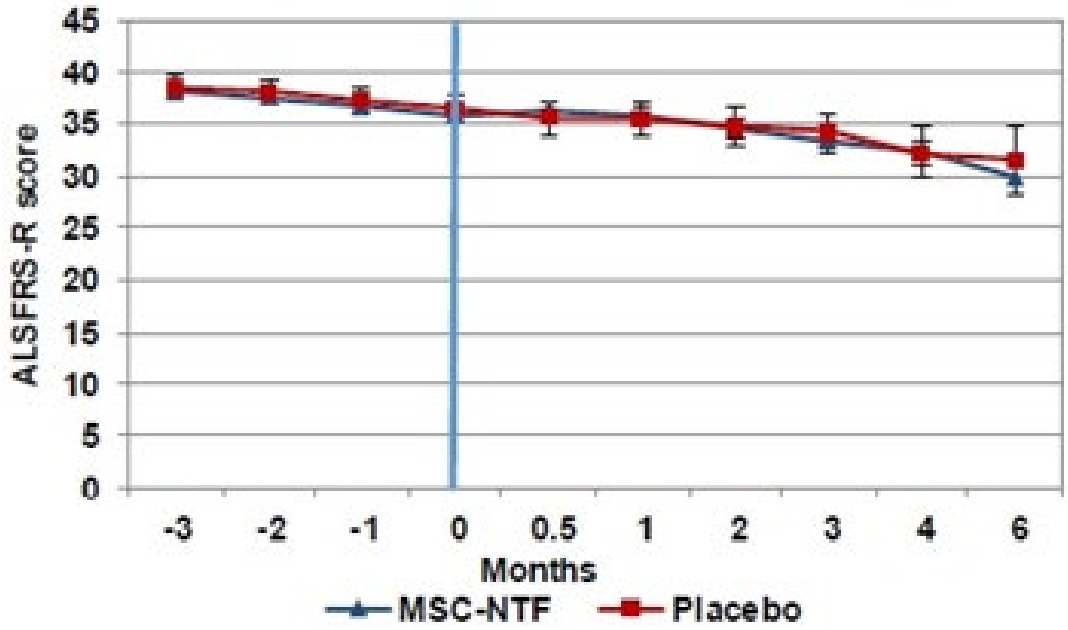
medication (four subjects, 8.3%). Post-treatment use of riluzole was reported in 19 (52.8%) and six (50.0%) subjects in the MSC-NTF and placebo groups, respectively.

Efficacy findings in the full analysis set, which included all 48 subjects, were negative and revealed no treatment benefit.

The pre-transplantation slopes were comparable between the two treatment groups. The means (SE) for the MSC-NTF treated group (n=36) and placebo (n=12) were -0.7 (0.1421) and -0.638 (0.2461), respectively. There was no significant difference (-0.2 [0.6], p=0.72) in post-pre-treatment slope changes between MSC-NTF and placebo groups.

As shown in Figure 15, ALSFRS-R total scores were comparable at baseline and throughout the 24 weeks of follow-up. ALSFRS-R total scores were 38.1 ± 3.5 (mean ± SD, observed data) and 38.5 ± 3.7 at screening, 35.9 ± 4.6 and 36.5 ± 4.5 at baseline/treatment, 30.0 ± 8.9 and 31.2 ± 9.2 at week 24 for MSC-NTF and placebo groups respectively. Changes from baseline to Week 24 were -5.9 ± 6.7 and -5.3 ± 6.6 for MSC-NTF and placebo groups, respectively.

**Figure 15. ALSFRS-R Score by Month, All Subjects**




Source: IND 15878
Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised.

There was no significant difference in SVC and muscle strength as measured by hand-held dynamometry.

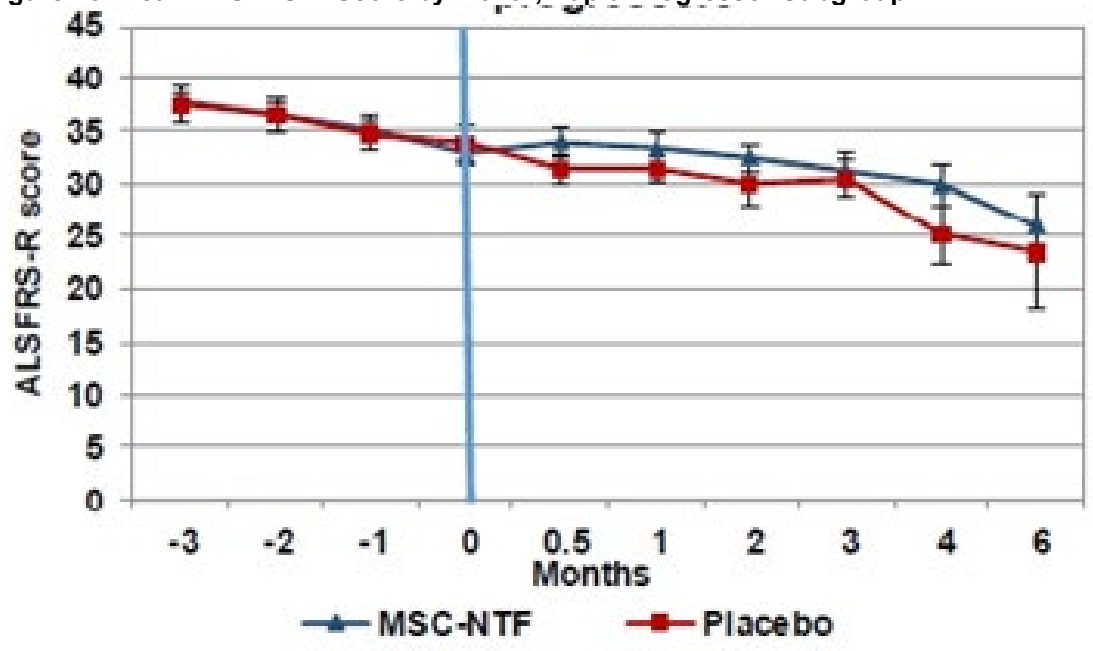
It is important to note that subjects in the Phase 2 study (BCT-001-US) had higher ALSFRS-R total scores at baseline than those enrolled in the Phase 3 study (BCT-002-US). Baseline scores of 35.9 ± 4.6 and 36.5 ± 4.5 were similar to those of subjects enrolled in the sodium phenylbutyrate/taurursodiol and tofersen studies. Per the Applicant's definitions, subjects enrolled in the Phase 2 study were less likely to be affected by floor effects. Therefore, the lack of treatment benefit in the Phase 2 study was unlikely due to floor effects.

The Applicant conducted exploratory subgroup analysis of "rapid progressors" versus "slow progressors." The Applicant defined "rapid progressors" as subjects with ≥2 points decline from screening to baseline (~3 months) in the ALSFRS-R total score; correspondingly, "slow progressors" were

defined as subjects with <2 points decline from screening to baseline in ALSFRS-R total score. As FDA stated in the November 18, 2019, Type C Meeting Summary:

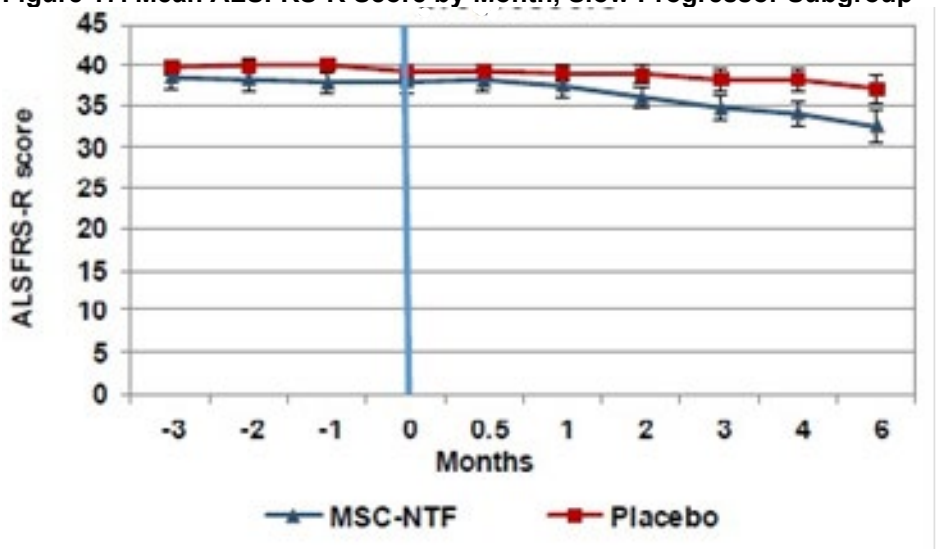
*"We interpret your Phase 2 data as evidence that your product is not effective in the treatment of ALS. Your proposal that your Phase 2 data suggest benefit for the 'rapid progressors' is most likely over-interpretation of your subgroup analyses. In subgroup analyses, the results for the 'slow progressors' could be interpreted to suggest that your product is harmful to some patients with ALS. However, such subgroup results, for both the 'rapid progressors' and the 'slow progressors', are most likely spurious and misleading, as is often the case for such subgroup analyses. We note that it is not clear why a product that you propose to have neuroprotective and immunomodulatory effects would be beneficial for some patients with ALS and harmful to other patients with ALS. Due to their inconsistency (i.e., opposite effects in 'rapid progressors' versus 'slow progressors'), and the unclear biological plausibility for such inconsistency, your subgroup results do not support that your product has any meaningful activity in the treatment of ALS"* (Figure 16 and Figure 17).

**Figure 16. Mean ALSFRS-R Score by Month, Rapid Progressor Subgroup**




Source: IND 15878
Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised.

**Figure 17. Mean ALSFRS-R Score by Month, Slow Progressor Subgroup**




Source: IND 15878
Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised.

Despite FDA's consistent concern about the definition of "rapid progressors," and the exploratory nature of the subgroup findings, the Applicant decided to enroll only "rapid progressors" in the Phase 3 study. For that study, the Applicant modified the definition of a "rapid progressor" to be subjects who experienced at least a 1.0-point decline in ALSFRS-R per month, on average, during the 3-month pre-treatment period.

## 4.4 Intermediate-Size Expanded Access Protocol

The primary objective of the ongoing intermediate-size EAP (BCT-003-US) is to provide MSC-NTF treatment for up to 12 subjects who completed the Phase 3, BCT-002-US study. The EAP consists of two treatment periods. In Treatment Period 1, subjects receive up to three intrathecal administrations of MSC-NTF every 8 weeks and will be followed for 12 weeks after the third dose MSC-NTF. In Treatment Period 2, subjects receive up to three intrathecal administrations of MSC-NTF every 8 weeks again. Following the last MSC-NTF dose, subjects will be followed for three months with monthly visits at which the ALSFRS-R and safety assessments will be conducted.

As of the data lock point of January 31, 2022, a total of 10 subjects who completed the Phase 3 study had been enrolled in the EAP, and all 10 subjects had completed EAP Treatment Period 1. Dosing for Treatment Period 2 was planned to begin on March 24, 2022. Of the 10 subjects, six received MSC-NTF and four received placebo in the Phase 3 study. Eight subjects who completed all three MSC-NTF treatments and the follow-up visit of Period 1 plan to continue in Period 2 of the EAP to receive three additional MSC-NTF treatments.

The EAP is not intended to provide evidence of effectiveness. All evaluations were assessed in an open-label, uncontrolled manner. The sample size of 10 is small. The interval between completion of the Phase 3 study and the first dose in EAP varied widely, ranging between 7 and 26 months. All 10 subjects treated in EAP had limb-onset ALS, as well as higher baseline ALSFRS-R entering the Phase 3 study (average 35.8± 2.9). It is challenging to interpret the findings in EAP subjects. As stated in the FDA guidance for industry *Expanded Access to Investigational Drugs for Treatment Use — Questions and*

*Answers*,[20] (October 2017) "Expanded access INDs and protocols are generally not designed to determine the efficacy of a drug; however, the expanded access regulations do not prohibit the collection of such data. Because expanded access INDs or protocols typically involve uncontrolled exposures (with limited data collection), it is unlikely that an expanded access IND or protocol would yield efficacy information that would be useful to FDA in considering a drug's effectiveness."

# 5. Biomarker Analysis

As discussed above, MSC-NTF is a bone marrow-derived, autologous cell therapy consisting of MSCs induced to secrete multiple neurotrophic factors (NTFs). After intrathecal administration, MSC-NTF is proposed to simultaneously deliver multiple NTFs into the CSF. Additional possible mechanisms could include secretion of immunomodulatory cytokines.

## 5.1 Overview of Biomarker Profiles After Administration of MSC-NTF

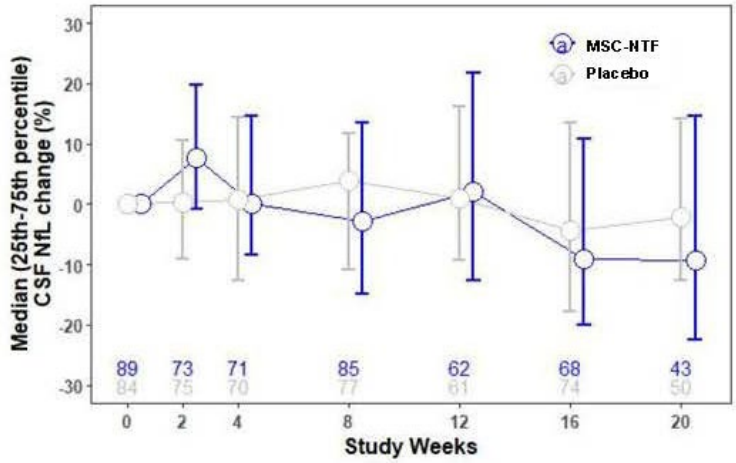
Subjects in Study BCT-002-US received either MSC-NTF or placebo, via intrathecal administration at Weeks 0, 8, and 16. Subjects in the MSC-NTF group received 100-125 × $10^6$ MSC-NTF cells at each intrathecal administration. To evaluate the relationship between biomarkers and treatment with MSC-NTF, CSF samples were collected at baseline and at study Weeks 2, 4, 8, 12, 16, and 20 post the first treatment. The Applicant evaluated a panel of 45 biomarkers, in four categories: neuroinflammation (anti-inflammatory and pro-inflammatory), neurodegeneration, neuroprotection, and other. Of note, a large amount of data was missing (~ 50%) at the Week 20 time point for all biomarkers (Appendix VIII).

Based on numerous exploratory analyses of all the biomarkers, the Applicant chose to emphasize the following in the BLA submission: NfL (neurodegeneration), galectin-1 (neuroprotection), LAP of TGF-beta-1 (referring to this as LAP moving forward) (anti-inflammatory), MCP-1 (pro-inflammatory), and VEGF-A (neuroprotection). Figure 18 shows the longitudinal changes in these CSF biomarkers in Study BCT-002-US.

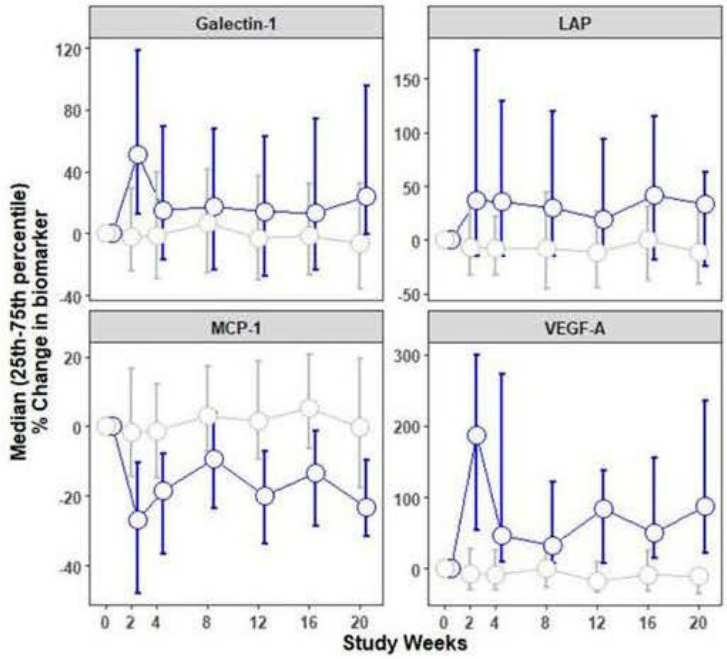
[20] FDA, 2017, *Guidance for Industry: Expanded Access to Investigational Drugs for Treatment Use —Questions and Answers*, https://www.fda.gov/regulatory-information/search-fda-guidance-documents/expanded-access-investigational-drugs-treatment-use-questions-and-answers

**Figure 18. Longitudinal Changes in Selected CSF Biomarkers**

**A. NfL**




**B. Galectin-1, LAP, MCP-1, and VEGF-A**




Source: FDA
Note: MSC-NTF: blue color; Placebo: grey color
Abbreviation: CSF, cerebrospinal fluid; NfL, neurofilament light chain.

NfL is a neurofilament protein that is highly expressed in myelinated axons. Elevated levels of NfL in CSF and blood are a consequence of axonal damage, and are observed in a variety of neurological disorders, including ALS. At Week 20, CSF NfL was reduced by 9.41% from baseline in MSC-NTF-treated subjects, compared to a 0.53% reduction in placebo-treated subjects (difference [ratio] for MSC-NTF to placebo: 8.88%; nominal p-value: 0.191).

## 5.2 Relationships Between Biomarkers and Clinical Efficacy Outcome

To evaluate the relationships between the selected CSF biomarkers and clinical efficacy outcomes, the Applicant conducted numerous exploratory analyses, including multiple post-hoc analyses. FDA independently performed correlation analyses of biomarker changes from baseline at Week 20, and change in ALSFRS-R total score from baseline to Week 28. (The Applicant did not collect biomarker samples beyond Week 20.) FDA's analyses of the biomarkers are described below.

Figure 19 shows the subject-level relationship between the percent change from baseline at Week 20 for the CSF biomarkers and change in ALSFRS-R total score from baseline to Week 28. These analyses utilized placebo and MSC-NTF treatment data from Study BCT-002-US. Correlation coefficients from Pearson and Spearman analyses are provided, together with the nominal p-values. In the mITT population (N=83), greater reduction of CSF NfL levels from baseline at Week 20 was seen in subjects with poorer clinical outcomes (gauged by change in ALSFRS-R total score from baseline to Week 28); since NfL levels are known to increase with increasing axonal damage, this result is the opposite of what would be expected (Figure 19A). Similar trends in correlation were also observed at an earlier time point, for change in NfL from baseline to Week 16 and change in ALSFRS-R total score from baseline to Week 28. Additional exploratory subgroup analysis showed the same trend in subjects with baseline ALSFRS-R >25 (i.e., the "no floor effect" subgroup, per Definition 1) (N=67) (Figure 19A). These findings could be due to 50% of missing NfL data at Week 20 and relatively overall small changes in NfL in MSC-NTF group.

No evident association was observed between other biomarkers percent change from baseline at Week 20 and ALSFRS-R changes from baseline to Week 28 (Figure 19B).

For a post-hoc subgroup of subjects whose ALSFRS-R item level had no value 0 at baseline (i.e., "no floor effect" subgroup, per Definition 3) (N=22), the Applicant further conducted post-hoc analyses using causal inference framework to study the relationship between change in CSF NfL levels and change in ALSFRS-R. Due to a large amount of missing data for the biomarker, a substantial amount of data imputation in the biomarker was done and all the models used were post-hoc, so can be selected to yield a result more favorable to MSC-NTF. Because these are post-hoc exploratory analyses in a subgroup of subjects, the results are likely biased and not interpretable. Thus, we believe there was no clear association observed between CSF NfL levels change from baseline at Week 20 and ALSFRS-R changes from baseline to Week 28 (Figure 19A).

The Applicant's NfL results are in contrast to the results obtained with tofersen. Tofersen is an antisense oligonucleotide designed to bind and degrade SOD1 messenger ribonucleic acid to reduce synthesis of the SOD1 protein. In a randomized, double-blind, placebo-controlled study, substantial and sustained reduction in plasma NfL was observed at Week 28 in the tofersen group compared to the placebo group (67% difference in geometric mean ratios for tofersen to placebo, nominal $p<0.0001$); total CSF SOD1, an indirect measure of target engagement, was reduced at Week 28 in the tofersen group compared to the placebo group (34% difference in geometric mean rations for tofersen to placebo, nominal $p<0.0001$); correlation analysis showed that reduction in plasma NfL was associated with less decline of ALSFRS-R from baseline.

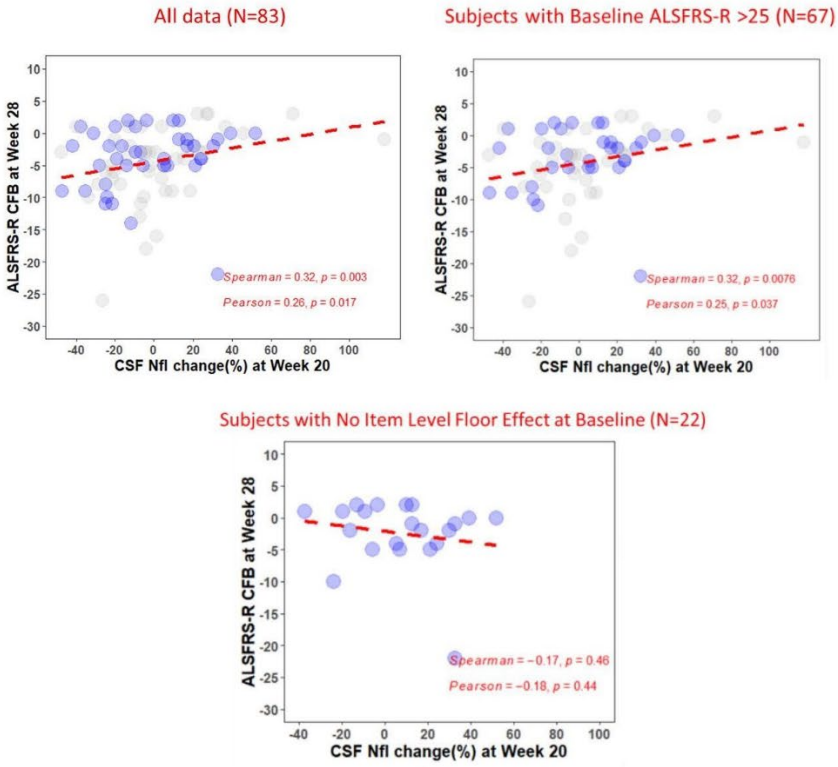
Additional data from an open-label extension study suggest reductions in NfL were seen in subjects previously receiving placebo who initiated tofersen in the open-label extension study, similar to the reductions seen in subjects treated with tofersen in the randomized, placebo-controlled study. Earlier

initiation of tofersen compared to placebo/delayed initiation of tofersen was associated with trends for reduction in decline on ALSFRS-R, SVC percent-predicted, and hand-held dynamometry megascore. Earlier initiation of tofersen was also associated with a trend toward reduction of the risk of death or permanent ventilation.[21] In contrast, the reduction in NfL after MSC-NTF was transient and of much smaller magnitude, and higher reduction in NfL after MSC-NTF was associated with worse clinical efficacy outcome. Therefore, the NfL data from the MSC-NTF Phase 3 study do not support the use of NfL as either a surrogate endpoint reasonably likely to predict clinical benefit for accelerated approval or as a biomarker to provide supportive evidence of effectiveness for traditional approval of MSC-NTF.

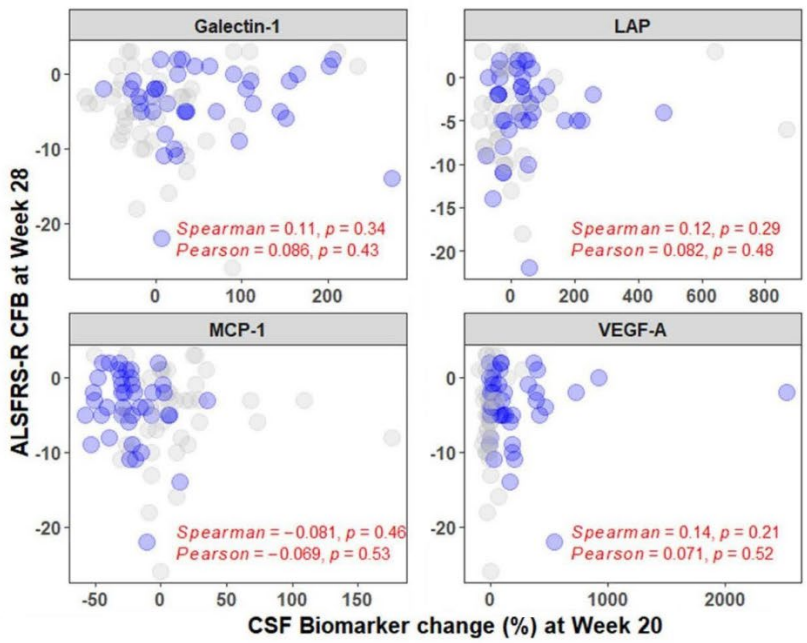
[21] Biogen, 2023, tofersen prescribing information, https://www.accessdata.fda.gov/drugsatfda_docs/label/2023/215887s000lbl.pdf.

**Figure 19. Subject-Level Relationship Between CSF Biomarkers Percent Change from Baseline at Week 20 and ALSFRS-R Changes from Baseline to Week 28**

**A. NfL**




**B. Galectin-1, LAP, MCP-1, and VEGF-A**




Source: FDA
Note: MSC-NTF: blue color; Placebo: grey color
Abbreviations: ALSFRS-R, ALS Functional Rating Scale–Revised; CFB, change from baseline; CSF, cerebrospinal fluid.

## 5.3 FDA's Concerns Regarding Biomarker Analyses

In Study BCT-002-US, the biomarker analyses were limited by the large amount of missing data. Biomarker data were only collected up to Week 20, but the efficacy data were collected up to Week 28. At Week 20, the key biomarkers that the Applicant identified, NfL, galectin-1, MCP-1, VEGF-A, and LAP, had up to approximately 50% missing data. In general, this degree of missing data compromises the validity of the analyses and could lead to over-estimation of the correlations between the biomarkers and efficacy endpoints. This missing data problem was further exacerbated when those post-hoc subgroup analyses were conducted based on different "floor effect" hypothesis.

Although the Applicant added the biomarker addendum to the statistical analysis plan before the data were unblinded, numerous biomarker analyses were proposed without multiplicity adjustment or formal hypothesis testing. The results from those biomarker analyses can only be considered as exploratory because there was no overall Type I error rate control, and any nominal "statistical significance" claim (nominal $p \leq 0.05$) could be due to chance alone.

Additionally, the applicant conducted multiple post-hoc analyses after the data were unblinded. These post-hoc analyses could be biased as the data are unblinded and analyses can be made to produce a more favorable result. Thus, post-hoc analyses in general have a high chance of false positive findings.

In summary, FDA does not believe there is sufficient evidence to support that any of the assessed biomarkers is reasonably likely to predict clinical benefit. Considering the potential mechanism of action of MSC-NTF, which may involve multiple pathways, it is challenging to use biomarker data to support effectiveness of MSC-NTF based on exploratory analyses of multiple biomarkers. There were also large amounts of missing data. In the case of NfL, which is released into the CSF by damaged or degenerating axons, higher reduction from baseline at Week 20 of CSF NfL levels were seen in subjects with poorer efficacy outcome (measured by ALSFRS-R score changes from baseline at Week 28), the opposite of what would be expected. These findings could be due to 50% of missing NfL data at Week 20 and relatively overall small changes in NfL in MSC-NTF group. Either way in the setting of negative phase 3 study results, the findings related to NfL do not appear to provide direct evidence on treatment effect through changes in NfL.

# 6. Safety Issues

Although safety is not the focus of this advisory committee meeting, the following data are provided for completeness.

## 6.1 Overview of Safety

### 6.1.1 Sources of Data for Safety

Study BCT-002-US was the only study that evaluated the safety of MSC-NTF administered via the intended intrathecal dosing regimen (three intrathecal administrations of MSC-NTF at 100-125 × $10^6$ cells, 8 weeks apart). In the earlier, single-arm studies (MSC-NTF-001-IL and MSC-NTF-002-IL), MSC-NTF was administered once at a different dose via intramuscular, intrathecal, or both intramuscular and intrathecal routes. In the Phase 2 study (BCT-001-US), MSC-NTF was administered once via combined intramuscular and intrathecal routes (single intrathecal administration of MSC NTF (~100-125 × $10^6$ cells) and intramuscular administration of MSC-NTF for a total of 48 × $10^6$ cells). Therefore, the safety review in this briefing document is primarily focused on safety data collected during the Phase 3 Study BCT-002-US.

In Study BCT-002-US, subjects in the MSC-NTF group received 100-125 × $10^6$ cells in DMEM per treatment and subjects in the placebo group received only DMEM. In the MSC-NTF group, 95 subjects received at least one treatment. Of these, 92 (96.8%) subjects received two treatments and 73 (76.8%) received all three treatments. In the placebo group, 94 subjects received at least one treatment. Of these, 84 (89.4%) received two treatments and 80 (85.1%) received all three treatments.

### 6.1.2 Deaths

There were 14 deaths in the mITT population during or shortly after the 28-week study follow-up period. Thirteen subjects (10 in MSC-NTF, 3 in placebo group) died within the study period (i.e., 28 weeks + 5 days). (See Section 4.2.3.2 Survival Was Worse in the MSC-NTF Group for more details.) In addition, two subjects (both randomized to the placebo group) died prior to receiving any assigned treatment. The most common cause of death was respiratory failure. One subject in MSC-NTF group died of a massive saddle embolism. One subject in MSC-NTF group elected voluntary euthanasia. One death in placebo group was due to drowning.

### 6.1.3 Safety Overview of Study BCT-002-US

A total of 1,892 TEAEs were reported in 186 subjects (98.4%), of which 1,020 events occurred in 94 subjects (98.9%) in the MSC-NTF treatment group and 872 events occurred in 92 subjects (97.9%) in the placebo group. Most TEAEs were reported as mild or moderate in severity (Table 4).

One subject in the MSC-NTF group and three subjects in the placebo group discontinued the study due to an adverse event.

No dose adjustments were made during the study.

Twenty-three (24.2%) subjects in the MSC-NTF group and 17 (18.1%) subjects in the placebo group had at least one treatment emergent serious adverse event (SAE). SAEs requiring hospitalization were comparable between the MSC-NTF and placebo groups. However, the MSC-NTF group had higher numbers of life-threatening SAEs (7, 7.4%) and SAEs with fatal outcomes (10, 10.5%) compared with the placebo group (1, 1.1% and 4, 4.3% for life-threatening SAEs and fatal SAEs respectively).

**Table 4. Summary of Treatment-Emergent Adverse Events**

| Adverse Events | MSC-NTF N=95 n (%) | Placebo N=94 n (%) | Relative Risk (95% CI) |
|---|---|---|---|
| SAE | 23 (24.2) | 17 (18.1) | 1.34 (0.77, 2.34) |
| SAEs with fatal outcome | 10 (10.5) | 4 (4.3) | 2.47 (0.80, 7.61) |
| Life-threatening SAEs | 7 (7.4) | 1 (1.1) | 6.93 (0.87, 55.21) |
| SAEs requiring hospitalization | 15 (15.8) | 16 (17.0) | 0.93 (0.49, 1.77) |
| SAEs resulting in substantial disruption of normal life functions | 0 (0.0) | 1 (1.1) | — |
| Congenital anomaly or birth defect | 0 (0.0) | 0 (0.0) | — |
| Other | 0 (0.0) | 0 (0.0) | — |
| AE leading to permanent discontinuation of study drug | 1 (1.1) | 3 (3.2) | 0.33 (0.03, 3.11) |
| Any AE | 94 (98.9) | 92 (97.9) | 1.01 (0.97, 1.05) |
| Severe | 29 (30.5) | 19 (20.2) | 1.51 (0.91, 2.50) |
| Moderate | 77 (81.1) | 67 (71.3) | 1.14 (0.97, 1.34) |
| Mild | 89 (93.7) | 89 (94.7) | 0.99 (0.92, 1.06) |

Source: FDA
Note: Frequencies tabulated based on actual treatment arm. AEs subset to treatment-emergent events. Percentages calculated with treatment arm totals as denominator. All counts represent unique subjects within each subgroup.
Abbreviations: AE, adverse event; SAE, serious adverse event.

TEAEs affecting at least 10% of the subjects are listed in Appendix V. The most common TEAEs affecting at least 20% of subjects were procedural pain, headache, back pain, procedural headache, fall, and post lumbar puncture syndrome.

The increase in the frequency of pain (back pain, musculoskeletal pain, and coccydynia), muscle spasms and dysphagia (Appendix VI) may affect the quality of life of the already vulnerable patients with ALS adversely.

Respiratory failure and dysphagia were the most reported serious TEAEs, consistent with the underlying disease and its complications. A total of 10 subjects, seven (7.4%) in the MSC-NTF group and three (3.2%) in the placebo group, had respiratory failure[22]. A total of five subjects, three (3.2%) in the MSC-NTF group and two (2.1%) in the placebo group, reported dysphagia (Appendix VII).

### 6.2 Safety Summary

(1) The higher incidence of deaths in the MSC-NTF group which indicates lack of survival benefit of MSC-NTF and warrants further investigation.

(2) There appears to be a higher incidence of respiratory failure and dysphagia in the MSC-NTF group.

(3) There appears to be a higher incidence of pain (e.g., coccydynia and back pain) in the MSC-NTF group.

## 7. Draft Voting Question for the Committee

(1) PLEASE VOTE:

Has substantial evidence of effectiveness meeting the approval standard been demonstrated by the evidence presented?

(a) Yes.

(b) No.

(2) If the answer to the above question is no, please discuss potential designs for a trial to demonstrate substantial evidence of effectiveness for MSC-NTF.

[22] Respiratory failure includes respiratory failure, respiratory distress, respiratory arrest.

# Appendices

## Appendix I. Summary of Regulatory History of BLA 125782

**Table 5. Summary of Regulatory History of BLA 125782**

| Date | Description |
|---|---|
| February 4, 2011 | Orphan Drug designation granted |
| December 20, 2013 | Original IND submission |
| October 2, 2014 | Fast Track designation granted |
| November 21, 2016 | Type B[1] End-of-Phase 2 Meeting: initial discussions regarding Applicant's plans for Phase 3 study<br>• FDA stated that Applicant's "proposed primary efficacy endpoint (i.e., the proportion of subjects whose disease is stabilized or improved as measured by a ≥1.5 point/month improvement in post-treatment slope versus pretreatment slope in ALSFRS-R score at 24 weeks from the first treatment) may potentially be acceptable for a Phase 3 study intended to support a market application," but that FDA "strongly [recommended] that [the Applicant] consider a primary efficacy endpoint whose clinical meaningfulness is easier to interpret (e.g., survival [death or tracheostomy]; CAFS)." FDA also provided CMC advice regarding concerns about variabilities noted in the manufacturing process, product comparability with the proposal to add new manufacturing sites, manufacturing scale, product specifications, material qualification, and the injection device. |
| August 7, 2017 | Type C[1] meeting to discuss Applicant's proposed Phase 3 protocol<br>• FDA expressed concerns regarding the clinical meaningfulness of defining a "responder" as a subject showing improvement of ≥1.25 in the slope of the least squares regression line for ALSFRS-R over time after treatment with MSC-NTF or placebo, compared with the slope prior to initiating treatment.<br>• FDA strongly recommended that the Applicant instead use as the primary efficacy endpoint for the Phase 3 study survival/tracheostomy or the CAFS.<br>• FDA recommended that the Applicant obtain an SPA[2] for the Phase 3 trial. |
| April 12, 2018 | RMAT designation denied |
| November 18, 2019 | Type C[1] meeting (Face-to-Face)<br>• The Applicant's Phase 2 study failed to show a statistically significant benefit of treatment with MSC-NTF compared with placebo.<br>• Although the Applicant's subgroup analysis seems to suggest potential benefit for "rapid progressors," that analysis correspondingly suggests potential harm to "slow progressors." FDA feels that these inconsistent effects most likely are spurious.<br>• The Applicant's Phase 3 study was underway at the time of this meeting. FDA again expressed concern regarding the study's primary efficacy endpoint, which uses ALSFRS-R linear regression slope to determine response to treatment.<br>• The Applicant expressed openness to modifying the primary and secondary efficacy endpoints. At that point, however, because of the importance of maintaining the integrity of the ongoing trial, FDA recommended that the Applicant not change the existing primary and secondary efficacy endpoints. FDA reasoned that since the Applicant was obtaining survival and joint-rank data, the study results would still be interpretable. FDA offered to review the Phase 3 study data prior to formal regulatory submission.<br>• FDA encouraged the Applicant to submit an SPA[2] request for a future Phase 3 study, prior to initiating that study. |

| Date | Description |
|---|---|
| February 5, 2020 | CBER Informal Dispute Resolution meeting with the Applicant to discuss outstanding issues, including efficacy endpoints, and analysis of the Applicant's ongoing Phase 3 study<br>• FDA disagrees regarding the primary efficacy endpoint but noted that in the study the Applicant was collecting data that the Agency considers critical to assess the efficacy of the product. FDA committed to reviewing the data, once the study is completed, to determine if there is a regulatory path forward that could potentially lead to approval.<br>• To maximally expedite further evaluation of this product for patients with ALS, FDA expressed willingness to explore ways to review the data prior to formal regulatory submission. |
| August 5, 2020 | Expanded access protocol (intermediate-size) submitted |
| December 16, 2020 | Type C[1] meeting to discuss the Applicant's preparation for commercial manufacturing, including proposed major manufacturing changes, critical materials, and acceptability of the Applicant's control strategy.<br>• FDA expressed concerns about the proposed changes, product comparability, and elements of the control strategy. |
| February 16, 2021 | Informal teleconference with CBER leadership and Applicant to discuss development of MSC-NTF for patients with ALS<br>• FDA does not consider the Phase 2 study results to provide evidence of efficacy of MSC-NTF for treatment of ALS.<br>• The Phase 3 study did not demonstrate a statistically significant benefit over placebo for the overall study population. FDA interprets the Phase 3 study results as a negative trial.<br>• The Applicant's suggested "floor effect" did not appear to be consistent across the study population: Subjects treated with placebo did not show the same "floor effect" proposed for subjects who received MSC-NTF. Since the study population was randomized, and the MSC-NTF and placebo groups were well-matched, a similar "floor effect" would be expected to be present in both groups.<br>• FDA remains concerned about the larger number of deaths among subjects treated with MSC-NTF compared to that among subjects treated with placebo.<br>• FDA does not consider submission of a BLA to be appropriate at this time. The available data do not meet the standard required for approval by either the traditional or the accelerated pathway. While the Applicant has the option of submitting a BLA with the existing data, the Agency may refuse to file it. In that event, the Applicant may choose to File Over Protest. The BLA will be reviewed, and an Advisory Committee meeting may be convened for additional discussion.<br>• If the Applicant wishes to continue pursuing development of this product for ALS, FDA recommends that the Applicant conduct another Phase 3 study, this time incorporating Agency input regarding the study design. |
| March 2, 2022 | FDA Public Statement on Amyotrophic Lateral Sclerosis Product Development [excerpt]:<br>*FDA knows that ALS patients, their families, and others in the ALS community have been closely watching the development of BrainStorm Cell Therapeutics, Inc.'s NurOwn [MSC-NTF] therapy. From data that have been communicated to the ALS community, there was a lot of hope that this product could provide at least a modest breakthrough in the management of ALS, if not something more substantial. Although FDA generally cannot provide confidential information about unapproved products, given the tremendous public interest in this product, we have concluded that it is important to provide high-level information about the status of the NurOwn development program.*<br>*With the recent completion of a randomized Phase 3 controlled clinical trial comparing NurOwn to placebo, it has become clear that data do not support the proposed clinical benefit of this therapy. Data indicated that none of the primary or secondary endpoints were met in the group of patients who were randomized.* |

| Date | Description |
|---|---|
| September 9, 2022 | Applicant submitted BLA 125782 |
| November 8, 2022 | FDA issued RTF letter to the Applicant |
| January 11, 2023 | Type A[1] meeting to discuss deficiencies identified in the RTF letter<br>• The Applicant chose not to discuss the Chemistry, Manufacturing, and Controls deficiencies. Discussion therefore covered only clinical deficiencies.<br>• FDA recommended the Applicant obtain all necessary information to address the deficiencies, and then submit a new BLA. The new BLA should include results of another adequate and well-controlled Phase 3 study to provide substantial evidence of effectiveness of MSC-NTF.<br>• FDA stated that notwithstanding the Agency's recommendation, the Applicant does have the option of requesting that the present BLA be filed over protest. |
| February 7, 2023 | Applicant requested File Over Protest |
| February 7, 2023 | FDA filed BLA 125782 over protest |

Source: FDA

1. Three types of formal meetings are held between FDA staff and Sponsors/Applicants developing new treatments: Type A, Type B, and Type C. The meeting types and goal dates were established under the Prescription Drug User Fee Act (PDUFA). (1) Type A meetings are held to help advance a stalled product development program and take place within 30 days of FDA receipt of the written meeting request. (2) Type B meetings occur within 60 days of FDA receipt of the written meeting request and are conducted at specific points in product development: to obtain FDA input prior to submission of investigational new drug application (Pre-IND); certain End-of-Phase 1 meetings; End-of-Phase 2 and pre-Phase 3 meetings; and prior to submission of a Biologics License Application (Pre-BLA). (3) Type C meetings refer to all other formal meetings regarding the development and review of an investigational product; Type C meetings are scheduled to occur within 75 days of FDA receipt of the written meeting request. Reference: Guidance for Industry *Formal Meetings Between the FDA and Sponsors or Applicants* (May 2009). https://www.fda.gov/media/72253/download

2. SPA refers to the process by which the developer of an investigational product discusses with FDA the design and size of certain clinical studies or animal studies to determine if they adequately address scientific and regulatory requirements for a study that could support marketing approval. An SPA agreement indicates concurrence by FDA with the adequacy and acceptability of specific elements of the overall protocol design (e.g., entry criteria, dose selection, endpoints, and planned analyses) that are critical for ensuring that the trial can be considered an adequate and well-controlled study. Reference: Guidance for Industry *Special Protocol Assessment* (April 2018). https://www.fda.gov/media/97618/download

Abbreviations: ALS, amyotrophic lateral sclerosis; ALSFRS-R, ALS Functional Rating Scale-Revised; BLA, Biologics License Application; CAFS, Combined Assessment of Function and Survival; IND, investigational new drug application; RMAT, Regenerative Medicine Advanced Treatment; RTF, Refuse to File; SPA, Special Protocol Assessment.

## Appendix II. Information on the Biological Activity of VEGF

**Table 6. Information on the Biological Activity of VEGF**

| Referenced Source of Information | Relevant VEGF Concentration Information |
|---|---|
| Biological activity of commercially available recombinant human VEGF protein (rhVEGF), including WHO VEGF standards | $ED_{50}$ in the 1-8 ng/ml range |
| Supplier recommended concentrations for in vitro use of rhVEGF | 1-5 ng/ml (1000 – 5000 pg/ml) |
| Reported concentration of rhVEGF in research publications when used on neuronal cultures | 10-250 ng/ml (10,000-250,000 pg/ml) Lunn, et al. 2009[23]; Tolsa, et al., 2008[24]; Plitonen, et al., 2011[25] |
| Amount of rhVEGF used in a Phase 1 ALS study conducted by Van Damme, et al. 2020, providing daily intraventricular delivery of rhVEGF for 3+ months (Van Damme, et al. 2020[26] | Daily intracerebroventricular delivery of up to 2 ug/day of purified rhVEGF. The study failed to show clinical benefit and achieved maximum CSF concentration of around 200 pg/ml (representing <2% of the administered level), demonstrating the challenge of achieving high levels of VEGF in vivo. |

[23] Lunn JS, Sakowski SA, Kim B, Rosenberg AA, Feldman EL. Vascular endothelial growth factor prevents G93A-SOD1-induced motor neuron degeneration. Dev Neurobiol. 2009 Nov;69(13):871-84. doi: 10.1002/dneu.20747. PMID: 19672955; PMCID: PMC2853013

[24] Tolosa, Laia et al. "Vascular endothelial growth factor protects spinal cord motoneurons against glutamate-induced excitotoxicity via phosphatidylinositol 3-kinase." Journal of neurochemistry vol. 105,4 (2008): 1080-90. doi:10.1111/j.1471-4159.2007.05206.x

[25] Piltonen, M et al. "Vascular endothelial growth factor C acts as a neurotrophic factor for dopamine neurons in vitro and in vivo." Neuroscience vol. 192 (2011): 550-63. doi:10.1016/j.neuroscience.2011.06.084

[26] Van Damme, Philip et al. Brain communications vol. 2,2 fcaa160. 29 Sep. 2020, doi:10.1093/braincomms/fcaa160

## Appendix III. Primary and Key Secondary Efficacy Endpoints, mITT Population, Phase 3 Trial

**Table 7. Primary and Key Secondary Efficacy Endpoints, mITT Population, Phase 3 Trial**

| Endpoint/Statistic | MSC-NTF (N=95) | Placebo (N=94) |
|---|---|---|
| ALSFRS-R total score (0-48) ≥1.25 points improvement in slope | | |
| Yes, n (%) | 31 (32.6) | 26 (27.7) |
| Odds ratio (95% CI)[1] | 1.33 (0.632, 2.798) | — |
| p-value[2] | 0.45 | — |
| ALSFRS-R total score (0-48) ≥100% improvement in slope | | |
| Yes, n (%) | 13 (13.7) | 13 (13.8) |
| Odds ratio (95% CI)[1] | 0.998 (0.416, 2.395) | |
| ALSFRS-R change from baseline at Week 28 | | |
| LS mean (SE) | -5.52 (0.670) | -5.88 (0.665) |
| LS mean difference (SE) | 0.37 (0.926) | — |
| 95% CI | -1.47, 2.20 | — |
| CAFS score | | |
| LS mean (SE) | 96.5 (5.1) | 93.5 (5.1) |
| LS mean difference (SE) | 3.0 (7.3) | — |
| 95% CI | -11.4, 17.4 | — |
| Nominal p-value | 0.68 | — |
| SVC change from baseline at Week 28 | | |
| LS mean (SE) | -12.94 (1.795) | -11.55 (1.806) |
| LS mean difference (SE) | -1.39 (2.395) | — |
| 95% CI | -6.15, 3.38 | — |
| Survival | | |
| Number of subjects died (%) | 10 (10.5%) | 3 (3%) |
| Hazard ratio (95% CI)[2] | 3.3 (0.87, 12.66) | — |
| Nominal p-value[3] (log-rank) | 0.044 | — |

Source: FDA statistician

1. Odds ratio and 95% CI calculated from the logistic regression model. Odds ratio is the ratio of odds for being a Responder in MSC-NTF group and odds for being a Responder in Placebo group.
2. Calculated from the logistic regression model.
3. The term “nominal” indicating the p value lack of multiplicity protection and consequently lack of interpretability.

Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised; CAFS, Combined Assessment of Function and Survival; CI, confidence interval; LS least squares; mITT, modified intent-to-treat; SE, standard error; SVC, slow vital capacity.

## Appendix IV. Primary and Key Secondary Efficacy Endpoints, no Floor Effect Subgroup Analyses, Phase 3

**Table 8. Primary and Selected Key Secondary Efficacy Endpoints, no Floor Effect Subgroup**

| Efficacy Endpoints | Total Score Threshold (N=145) MSC-NTF | Total Score Threshold (N=145) Placebo | Item Level Threshold (N=159) MSC-NTF | Item Level Threshold (N=159) Placebo | Revised Item Level Threshold (N=106) MSC-NTF | Revised Item Level Threshold (N=106) Placebo |
|---|---|---|---|---|---|---|
| ALSFRS-R total score ≥1.25 points improvement in slope | | | | | | |
| Yes, n (%) | 25 (34.7%) | 15 (20.5%) | 28 (35.4%) | 18 (22.5%) | (40.8%) | 22.8% |
| Nominal p-value | 0.053 | — | 0.035 | — | 0.035 | — |
| ALSFRS-R total score ≥100% improvement in slope | | | | | | |
| Yes, n (%) | 12 (16.7%) | 10 (13.7%) | 13 (16.5%) | 11 (13.8%) | 12 (24.5%) | 8 (14%) |
| Nominal p-value | 0.804 | — | 0.681 | — | 0.282 | — |
| ALSFRS-R MMRM change from baseline at Week 28 | | | | | | |
| Missing (%) | 59 (13) | 57 (16) | 62 (17) | 61 (19) | 6 (12) | 12 (27) |
| LS mean (SE) | -4.82±0.73 | -5.98±0.72 | -4.60±0.69 | -6.08±0.68 | -2.68 (0.83) | -4.99 (0.78) |
| LS mean difference (SE) | 1.16±1.00 | — | 1.48±0.95 | — | 2.31 (1.1) | — |
| Nominal p-value | 0.25 | — | 0.125 | — | 0.04 | — |
| CAFS score | | | | | | |
| LS mean (SE) | 78.7 (4.2) | 67.4 (4.2) | 85.7 (4.5) | 74.4 (4.5) | 59.4 (4.0) | 48.4 (3.7) |
| LS mean difference (95% CI) | 11.3 (-0.5, 23.1) | — | 11.3 (-1.3, 23.9) | — | 11.0 (0.23, 21.9) | — |
| Nominal p-value | 0.06 | — | 0.08 | — | 0.045 | — |

Source: FDA statistician

Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised; CAFS, Combined Assessment of Function and Survival; CI, confidence interval; LS, least squares; MMRM, mixed effects repeated measures; SE, standard error.

**Table 9. Primary and Key Secondary Efficacy Endpoints by ALSFRS-R Total Score at Baseline (>25 versus ≤25)**

| Endpoint | ALSFRS-R Total Score at Baseline >25 (n=149) | ALSFRS-R Total Score at Baseline ≤25 (n=40) |
|---|---|---|
| ALSFRS-R total score (0-48) ≥1.25 points improvement in slope | | |
| Yes (%) | 34.7% vs. 20.5% | 26% vs. 52% |
| Odds ratio (95% CI) [1] | 2.529 (0.987, 6.478) | 0.178 (0.031, 1.017) |
| Nominal p-value[2] | 0.053 | 0.0522 |
| ALSFRS-R total score (0-48) ≥100% improvement in slope | | |
| Yes, (%) | 16.7% vs. 13.7% | 4% vs. 14% |
| Odds ratio (95% CI)[1] | 1.134 (0.422, 3.049) | 0.29 (0.018, 4.661) |
| Nominal p-value[2] | 0.804 | 0.38 |
| ALSFRS-R change from baseline at Week 28 | | |
| Missing (%) | | |
| LS mean (SE) | -4.8 (0.7) vs. -6.0 (0.7) | -8.6 (1.5) vs. -5.4 (1.5) |
| LS mean difference (SE) | 1.16 (1.0) | -3.2 (2.1) |
| 95% CI | (-0.8, 3.2) | (-7.4, 1.1) |
| Nominal p-value[2] | 0.25 | 0.14 |
| CAFS score | | |
| LS mean (SE) | 78.7 (4.2) vs. 67.4 (4.2) | 18.7 (2.3) vs. 26.7 (2.4) |
| LS mean difference (SE) | 11.3 (-0.5, 23.1) | -7 (-15.0, -1) |
| 95% CI | | |
| Nominal p-value[2] | 0.06 | 0.026 |
| SVC change from baseline at Week 28 | | |
| Missing (%) | | |
| LS mean (SE) | -11 (2.1) vs. -10.5 (2.2) | -18 (3.9) vs. -15 (3.4) |
| LS mean difference (SE) | -1.1 (-6.8, 4.6) | -3.4 (-13.2, 6.5) |
| 95% CI | | |
| Nominal p-value[2] | 0.7 | 0.5 |

Source: FDA statistician

1. Odds ratio and 95% CI calculated from the logistic regression model. Odds ratio is the ratio of odds for being a Responder in MSC-NTF group and odds for being a Responder in Placebo group.

2. The term “nominal” indicating the p value lack of multiplicity protection and consequently lack of interpretability.

Abbreviation: ALSFRS-R, ALS Functional Rating Scale–Revised; CAFS, Combined Assessment of Function and Survival; CI, confidence interval; LS, least squares; SE, standard error; SVC, slow vital capacity.

## Appendix V. Common Treatment Emergent Adverse Events Affecting ≥10% Subjects

**Table 10. Common Treatment Emergent Adverse Events Affecting ≥10% Subjects**

| Preferred Term | MSC-NTF (N=95) n (%) | Placebo (N=94) n (%) | Relative Risk RR (95% CI) |
|---|---|---|---|
| Procedural pain | 50 (52.6) | 34 (36.2) | 1.5 (1.0, 2.0) |
| Headache | 45 (47.4) | 32 (34.0) | 1.4 (1.0, 2.0) |
| Back pain | 42 (44.2) | 24 (25.5) | 1.7 (1.1, 2.6) |
| Procedural headache | 31 (32.6) | 30 (31.9) | 1.0 (0.7, 1.5) |
| Fall | 29 (30.5) | 34 (36.2) | 0.8 (0.6, 1.3) |
| Post lumbar puncture syndrome | 22 (23.2) | 29 (30.9) | 0.8 (0.5, 1.2) |
| Musculoskeletal pain[1] | 18 (18.9) | 9 (9.6) | 2.0 (0.9, 4.2) |
| Nausea | 16 (16.8) | 18 (19.1) | 0.9 (0.5, 1.6) |
| Pain in extremity | 16 (16.8) | 11 (11.7) | 1.4 (0.7, 2.9) |
| Post procedural complication | 16 (16.8) | 7 (7.4) | 2.3 (1.0, 5.2) |
| Muscle spasms[2] | 12 (12.6) | 6 (6.4) | 2.0 (0.8, 5.1) |
| Muscular weakness | 11 (11.6) | 12 (12.8) | 0.9 (0.4, 2.0) |
| Dysphagia | 11 (11.6) | 7 (7.4) | 1.6 (0.6, 3.8) |
| Coccydynia | 11 (11.6) | 1 (1.1) | 10.9 (1.4, 82.6) |
| Arthralgia | 10 (10.5) | 7 (7.4) | 1.4 (0.6, 3.6) |
| Laceration | 7 (7.4) | 11 (11.7) | 0.6 (0.3, 1.6) |
| Upper respiratory tract infection | 6 (6.3) | 12 (12.8) | 0.5 (0.2, 1.3) |

Source: FDA
1. Musculoskeletal pain includes musculoskeletal pain and myalgia.
2. Muscle spasms includes muscle spasms and muscle spasticity.
Abbreviation: TEAE, treatment-emergent adverse event.

## Appendix VI. Common Treatment-Emergent Adverse Events with Relative Risk >1.5

**Table 11. Common Treatment-Emergent Adverse Events With Relative Risk >1.5**

| Preferred Term | MSC-NTF (N=95) n/N(%) | Placebo (N=94) n/N (%) | Relative Risk RR (95% CI) |
|---|---|---|---|
| Coccydynia | 11/95 (11.6) | 1/94 (1.1) | 10.88 (1.43, 82.64) |
| Post procedural complication | 16/95 (16.8) | 7/94 (7.4) | 2.26 (0.98, 5.24) |
| Musculoskeletal pain[1] | 18/95 (18.9) | 9/94 (9.6) | 1.98 (0.94, 4.18) |
| Muscle spasms[2] | 12/95 (12.6) | 6/94 (6.4) | 1.98 (0.77, 5.05) |
| Back pain | 42/95 (44.2) | 24/94 (25.5) | 1.73 (1.15, 2.62) |
| Dysphagia | 11/95 (11.6) | 7/94 (7.4) | 1.55 (0.63, 3.84) |

Source: FDA
1. Musculoskeletal pain includes musculoskeletal pain and myalgia.
2. Muscle spasms includes muscle spasms and muscle spasticity.

## Appendix VII. Common Treatment Emergent Serious Adverse Events Affecting ≥2 Subjects

**Table 12. Common Treatment Emergent Serious Adverse Events Affecting ≥2 Subjects**

| Preferred Term | MSC-NTF (N=95) n (%) | Placebo (N=94) n (%) | Relative Risk RR (95% CI) |
|---|---|---|---|
| Respiratory failure[1] | 7 (7.4) | 3 (3.2) | 2.3 (0.6, 8.7) |
| Dysphagia | 3 (3.2) | 2 (2.1) | 1.5 (0.3, 8.7) |
| Pneumonia | 2 (2.1) | 2 (2.1) | 1.0 (0.1, 6.9) |
| Disease progression | 1 (1.1) | 2 (2.1) | 0.5 (0.0, 5.4) |

Source: FDA

[1] Respiratory failure includes preferred terms respiratory failure, respiratory arrest, respiratory distress

## Appendix VIII. Subject Number for Biomarker Analysis at Week 20

**Table 13. Subject Number for Biomarker Analysis at Week 20 (Study BCT-002-US, MSC-NTF-treatment Group)**

| Biomarker | Subject Number |
|---|---|
| **Neuroinflammation** | (b) (6) |
| Fetuin-A (anti-inflammatory) | |
| hsa-miR-146a-5p (anti-inflammatory) | |
| hsa-miR-146b-5p (anti-inflammatory) | |
| IL-37 (anti-inflammatory) | |
| LAP (anti-inflammatory) | |
| MSR-1 (anti-inflammatory) | |
| CHI3L1/YKL-40 (inflammatory) | |
| Chitotriosidase-1 (inflammatory) | |
| GFAP (inflammatory) | |
| ICAM-1 (inflammatory) | |
| IP-10 (inflammatory) | |
| MCP-1 (inflammatory) | |
| OPG (inflammatory) | |
| S100B (inflammatory) | |
| SDF-1a (inflammatory) | |
| TREM-2 (inflammatory) | |
| **Neurodegeneration** | |
| Caspase-3 | |
| DR6 | |
| hsa-miR-142-5p | |
| NfL | |
| pNFH | |
| Tau | |
| TWEAK | |
| UCH-L1 | |
| **Neuroprotection** | |
| BDNF | |
| Clusterin/ApoIJ | |
| G-CSF | |
| Galectin-1 | |
| GDF-15 | |
| HGF | |
| LIF | |
| NMNAT1 | |
| VEGF-A | |
| **Others** | |
| Follistatin | |
| hsa-miR-124-3p | |
| hsa-miR-126-3p | |
| hsa-miR-132-3p | |
| hsa-miR-199b-5p | |
| hsa-miR-19b-3p | |
| hsa-miR-206 | |
| hsa-miR-20a-5p | |
| hsa-miR-30b-5p | |
| hsa-miR-34a-5p | |
| hsa-miR-9-3p | |
| Osteopontin | |

Source: FDA

## Appendix IX. Percent Change From Baseline Key Biomarkers at Week 20

**Table 14. Percent Change From Baseline Key Biomarkers at Week 20 (Study BCT-002-US)**

| Biomarker | MSC-NTF | Placebo | p-value[1] Difference |
|---|---|---|---|
| NfL | -9.41% | -0.53% | 0.191 |
| MCP-1 | -23.27% | -2.24% | <0.001 |
| Galectin-1 | 12.74% | -7.73% | 0.116 |
| LAP (TGF-β1) | 12.23% | -21.51% | 0.028 |
| VEGF-A | 144.54% | -16.56% | <0.001 |

Source: FDA

1. Nominal p-values for the treatment difference in least square means from MMRM with treatment, visit, and treatment by visit interaction as fixed effects.

Abbreviations: LAP, latency-associated peptide; MCP-1, monocyte chemotactic factor-1; NfL, neurofilament light chain; VEGF, vascular endothelial growth factor.

## Appendix X. Issues with Subgroup Analyses

Descriptive and graphical subgroup analyses (based on baseline demographic and clinical characteristics) are routine in clinical trials and serve several functions. In the presence of an overall favorable result on efficacy across a whole population, subgroup analyses are used to evaluate the robustness and consistency of a treatment effect estimate. In the presence of an overall negative or null result on efficacy across a whole population, subgroup analyses can be used to generate hypotheses about why the result was negative and potentially identify subpopulations for more targeted follow-up trials.

Subgroup analyses, however, are not reliable for "overturning" a negative result on efficacy and concluding that an ineffective treatment works in a subpopulation. This is because of the statistical problems of multiple testing: The more comparisons that are made in a trial, the greater the chance that at least one will lead to a false positive result. This is true when evaluating pre-defined subgroup analyses and is particularly pernicious when the subgroup itself is defined on a post-hoc basis, because there is no practical limit to the number of subgroups that might have been so defined, and therefore no practical limit to the chance of a false positive conclusion on efficacy. This was well illustrated in the famous ISIS-2 trial, in which investigators demonstrated, as a pedagogical example of the problem of subgroups, that subjects with certain astrological signs (Gemini and Libra) were apparently significantly harmed by an intervention that was beneficial for patients born under luckier signs.[27]

[27] ISIS-2 (Second International Study of Infarct Survival) Collaborative Group. Randomised trial of intravenous streptokinase, oral aspirin, both, or neither among 17 187 cases of suspected acute myocardial infarction: ISIS-2. Lancet. 1988; ii: 349–360.