

FDA Briefing Document

**Peripheral and Central Nervous System Drugs
Advisory Committee Meeting**

April 25, 2016

**NDA 206488
Eteplirsen**

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 or Office. We have brought these issues to this Advisory Committee in order to gain the Committee's insights and opinions, and the background package may not include all issues relevant to the final regulatory recommendation and instead 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 be affected by issues not discussed at the advisory committee meeting.

Table of Contents

- I. Briefing Memorandum to the Committee
- II. Draft Points to Consider
- III. Clinical Team Leader Memorandum to the Committee
- IV. Statistical Review
- V. Summary of Clinical Pharmacology Findings

I. Briefing Memorandum to the Committee

MEMORANDUM

DATE: April 9, 2016

FROM: Billy Dunn, MD and Eric Bastings, MD, Division of Neurology Products
Ellis Unger, MD and Robert Temple, MD, Office of Drug Evaluation-I
Office of New Drugs, Center for Drug Evaluation and Research, FDA

TO: Members and Invited Guests of the Peripheral and Central Nervous Systems
Drugs Advisory Committee (PCNS AC)

SUBJECT: Briefing Memo for New Drug Application (NDA) 206488, for the use of eteplirsen
for the treatment of Duchenne muscular dystrophy in patients with mutations
amenable to exon 51 skipping

The Peripheral and Central Nervous Systems Drugs Advisory Committee will be meeting on April 25, 2016, to discuss the NDA for eteplirsen, submitted by Sarepta Therapeutics, Inc., for the treatment of certain patients with Duchenne muscular dystrophy (DMD). The Committee includes experts on DMD, neurology, clinical trial design, and biostatistics, as well as representatives from the DMD patient community. Sarepta is seeking accelerated approval for eteplirsen for patients with DMD who have a confirmed mutation of the dystrophin gene amenable to exon 51 skipping ($\approx 13\%$ of patients with DMD). In such patients, skipping of exon 51 might restore the reading frame of dystrophin, increase the production of dystrophin, and lead to a clinical benefit for patients.

The applicant undertook three studies: two small exploratory studies (Study 28 and Study 33) to assess eteplirsen's potential to increase dystrophin expression, and a 12-patient clinical study (Study 201/202) to further assess the extent to which eteplirsen increased expression of dystrophin protein, and to explore the potential clinical benefit. The designs and results of these studies have been reviewed in detail by a multidisciplinary review team led by Dr. Ronald Farkas (Cross-Disciplinary Team Leader). Included in this briefing package are an integrated summary review of the eteplirsen data by Dr. Farkas, a statistical review of Study 201/202 by Dr. Xiang Ling, and a summary of clinical pharmacology findings by Drs. Attul Bhattaram, Ta-Chen Wu, and Bart Rogers.

This Advisory Committee meeting was initially scheduled to take place on January 22, 2016, but had to be rescheduled because of a weather emergency. Since the initial FDA briefing materials were released, the applicant submitted additional information about clinical outcomes of

patients in Study 201/202, and also made public an addendum to their briefing materials in which the applicant describes what it calls “key inaccuracies” in the briefing document FDA released in advance of the original date for this Advisory Committee meeting. As will be discussed below, and in more detail in the Cross-Disciplinary Team Leader summary document, we do not agree with the applicant’s characterization of inaccuracies in the initial FDA briefing document.

As explained by the applicant, eteplirsen’s intended mechanism of action is removal of exon 51 of the pre-messenger ribonucleic acid (RNA), thereby restoring the messenger RNA “reading frame.” This shift would enable the production of a truncated form of the dystrophin protein. By increasing the quantity of an abnormal, but potentially functional, dystrophin protein, the objective is to slow or prevent the progression of DMD.

Pharmacodynamic and clinical effects, therefore, are potentially demonstrable at 3 levels: 1) expression of an altered messenger RNA in muscle (pharmacodynamic); 2) production of dystrophin protein in muscle (pharmacodynamic); and 3) improvement or preservation of muscle function (clinical).

As noted above, the applicant conducted 3 studies to assess the pharmacodynamic and/or clinical effects of eteplirsen. Study 33 was an exploratory phase 1 study in which small doses of eteplirsen (up to 0.9 mg) were injected directly into a foot muscle in seven patients with DMD. Study 28 was an exploratory study in which eteplirsen was administered intravenously once a week for 12 weeks at doses up to 20 mg/kg in 19 patients with DMD. Study 201/202 was the only concurrently controlled clinical trial conducted by Sarepta intended to assess a clinical endpoint. Study 201/202 began as a 24-week randomized placebo-controlled study (Study 201) comparing three groups of four patients each, treated weekly with intravenous eteplirsen 50 mg/kg, eteplirsen 30 mg/kg, or placebo (placebo patients were divided in two subgroups, one switched to eteplirsen 30 mg/kg at Week 24, and the other switched to eteplirsen 50 mg/kg at Week 24). The prospectively planned primary endpoint of Study 201 was an assessment of dystrophin in skeletal muscle. In Study 201, all 12 patients had a muscle biopsy at baseline (first biopsy) and Week 48 (third biopsy). In addition, patients had a second muscle biopsy either at Week 12 (50 mg/kg group) or Week 24 (30 mg/kg group). The randomized controlled phase (Study 201) was followed by an open-label extension phase in which all 12 patients received eteplirsen 30 mg/kg, weekly, by the intravenous route (Study 202). In Study 202, 11 of the 12 patients had a fourth muscle biopsy at Week 180 (~3.5 years).

1. Expression of the Expected mRNA in Muscle

The applicant evaluated the effect of eteplirsen on production of dystrophin messenger RNA in Study 33, Study 28, and Study 201/202.

Skipping of the mRNA exon was assessed using reverse transcriptase polymerase chain reaction (RT-PCR), a standard technique commonly used in molecular biology laboratories to detect RNA expression. The applicant notes that exon 51 skipping was confirmed by RT-PCR analysis in all patients treated with eteplirsen. PCR is a highly sensitive technique that can detect even a few copies of messenger RNA. Because even a minimal PCR signal is interpreted as “positive,” this biomarker provides little support of efficacy for eteplirsen; it does provide evidence that eteplirsen causes at least some degree of exon 51 skipping, as intended.

2. Production of Dystrophin Protein in Muscle

The applicant evaluated the effect of eteplirsen on dystrophin expression primarily in Study 201/202, but also in Studies 28 and 33. Production of dystrophin was assessed by two different methods: immunofluorescence (IF) and Western blot. In considering these two measures, it is important to note that Western blot is considered to be a quantitative method, whereas immunofluorescence is generally considered to be less quantitative, and is more often relied upon to show the localization of protein in tissue sections. The applicant used Western blot to quantify dystrophin protein directly. Immunofluorescence methods were used to distinguish “positive” muscle fibers, i.e., those with at least some degree of positivity, from “negative” muscle fibers in tissue biopsy sections, and the data were also analyzed based on the staining intensity of identified areas of tissue sections.

Immunofluorescence (IF)

The percentage of dystrophin-positive fibers in tissue obtained from muscle biopsies was the prospectively planned primary endpoint of Study 201.

Substantial increases in dystrophin in Study 201 were initially reported in a publication,¹ which stated the “...percentage of dystrophin-positive fibers was increased to 23% of normal; no increases were detected in placebo-treated patients ($p \leq 0.002$). Even greater increases occurred at week 48 (52% and 43% in the 30 and 50 mg/kg cohorts, respectively....)”.

¹ *Ann Neurol* 2013;74:637

FDA conducted an inspection of the facility where the data reported in the publication were generated. Significant methodological concerns were identified, which cast serious doubt on the reliability of assessments from the first three biopsies. In light of these concerns, FDA worked collaboratively with the applicant on methods for the reassessment of the images, as well as collection of additional data that could be more reliable. The goal of this effort was to help the applicant apply suitable, consistent, and objective methods for measuring dystrophin protein that would be amenable to independent verification for any future biopsies for patients in Study 201/202 and other planned studies.

Eleven (11) of the 12 patients in Study 201/202 consented to a fourth biopsy at Week 180 (3.5 years), with dystrophin levels to be compared to their archived pre-treatment tissue. Unfortunately, archived pre-treatment tissue was available for only 3 of the 11 patients. The applicant therefore supplemented these baseline samples with muscle tissue from 6 other untreated patients with DMD amenable to exon 51 skipping.

On re-analysis of the first three biopsies by the 3 blinded readers, the mean percent dystrophin-positive fibers for the 4 patients in the 30 mg/kg eteplirsen group was 14% at baseline, 27% at Week 24, and 23% at Week 48. For the 4 patients in the 50 mg/kg group, the mean percent dystrophin-positive fibers was 15% at baseline, 17% at Week 12, and 25% at Week 48 (Table 1).

Table 1: Study 201 immunofluorescence results for first three muscle biopsies (% positive fibers)

	Nationwide Children’s Hospital analysis				Re-analysis by 3 blinded readers				
	Baseline	Week 12	Week 24	Week 48	Baseline	Week 12	Week 24	Week 48	Week 180 (n=11)
30 mg/kg (n=4)	18		41	70	14		27	23	17
50 mg/kg (n=4)	11	12		54	15	17		25	
Placebo to 30 mg/kg (n=2)	24		24	58	10		10	9	
Placebo to 50 mg/kg (n=2)	7	7		49	11	9		10	

Biopsies up to Week 48 had methodological shortcomings, however, uncovered at the FDA facility inspection. Therefore, the results from the fourth (Week 180) biopsy are particularly important to the interpretation of the study results. For the 11 eteplirsen-treated patients who had a biopsy at Week 180, the three blinded veterinary pathologists reported a mean of 17% of dystrophin-positive fibers for the eteplirsen-treated patients, a level considerably lower than that reported for the first three biopsies in earlier reports.¹ Control patients had about 1% dystrophin-positive fibers.

Of note, in their January 2016 addendum, the applicant described as a “key inaccuracy” a statement by FDA that “the lack of an effect [on immunofluorescence results] with the higher dose group tends to undermine the finding in the lower dose group and the lack of even a positive trend at the earlier time point (with a higher dose) sheds doubt on the finding at a later time point.” The applicant argues that “duration of therapy was observed to be the critical variable when interpreting dystrophin levels. 12 weeks does not represent a clinically relevant duration of therapy.” However, the applicant also states in the briefing materials that, in Study 28, weekly treatments with eteplirsen for 12 weeks resulted in a “3-fold increase in the mean percentage of dystrophin-positive fibers.” Although these increases cannot be confirmed by FDA because of methodological issues, and were not confirmed in Study 201 (Table 1), it seems clear that the applicant considered increases in dystrophin-positive fibers after 12 weeks of treatment as possible, and it remains that the negative findings at a higher dose of eteplirsen at Week 12 weaken the findings at Week 24..

More importantly, we believe that analyses based on immunofluorescence overestimate the amount of dystrophin in tissue sections. This is because a muscle fiber can be considered “positive” if it exhibits any staining at all, even if the level of dystrophin is very low. Specifically, consider the following example: a microscopic field where 25% of fibers are counted as “positive,” but where their staining intensity is faint, perhaps 3% of normal brightness on average. Although some 25% of fibers are deemed to be “positive,” the overall dystrophin content could be estimated at $3\% \times 25\% = 0.75\%$. Thus, the review team does not consider “percent dystrophin-positive fibers” to be a meaningful way to estimate dystrophin content, and we believe the results reported by the applicant on this measure do not establish that a significant increase in dystrophin occurred in response to eteplirsen treatment.

Western Blot

The applicant provided a second line of evidence, Western blot analysis, to support the concept that eteplirsen increases dystrophin production in skeletal muscle. By Western blot, the most accurate quantitative method used by the applicant, the mean dystrophin level after ~3.5 years

of eteplirsen treatment was $0.93\% \pm 0.84\%$ of normal (mean \pm standard deviation). This 0.9% estimate is in stark contrast to the earlier reports of dystrophin-positive fibers, with reported increases to as great as 50% of normal,¹ levels that were based on methods we have determined were unreliable for accurate quantification. The more relevant and reliable quantitative dystrophin estimate of 0.9% of normal after 3.5 years of treatment is disappointing.

Table 2, adapted from the applicant’s submission, shows the anonymized adjudicated results for dystrophin quantification from the fourth biopsy as assessed by Western blot (percent of normal) and immunofluorescence (percent dystrophin-positive fibers) for the 11 patients who volunteered for muscle biopsies at Week 180. Overestimation by the immunofluorescence method is apparent.

Table 2: Applicant’s Quantification of Dystrophin by Western Blot and Immunofluorescence Analyses

Patient	Western Blot % of normal	Immunofluorescence % positive fibers
A	2.05	18.5
B	1.15	19.1
C	0.38	33.5
D	1.62	24.0
E	0.52	21.5
F	0.98	12.8
G	0	7.1
H	2.47	20.7
I	0.96	28.2
J	0	1.4
L	0.14	4.5

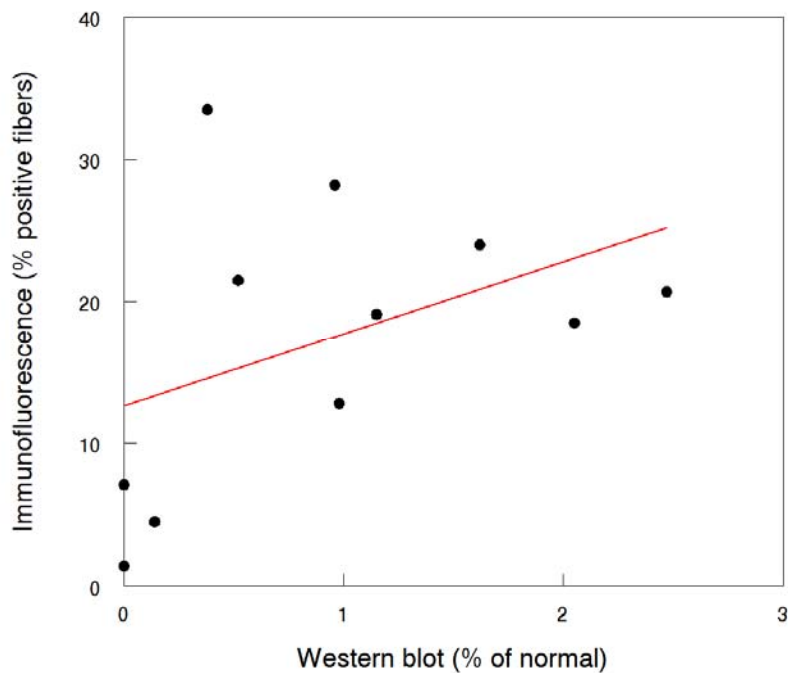
FDA had also suggested the applicant attempt to assess dystrophin levels at baseline, i.e., pre-treatment. The applicant reported a control (untreated) value of 0.08% dystrophin based on retained samples from the pre-treatment biopsy in 3 patients from Study 201/201 combined with data from six patients with DMD who were not enrolled in any study. The applicant suggests that these data support a conclusion of “an 11.6-fold increase in de novo dystrophin production was observed by Western blot relative to untreated controls.” (page 25 of their briefing book).

There are, however, some important limitations with respect to interpretation of the results of the untreated controls. 1) The reported mean value of 0.08% is well below the lower limit of detection of the applicant’s Western blot assay (0.25%); 2) Archived pre-treatment muscle

biopsy samples were available for re-analysis from only 3 patients in Study 201/202. Additional samples were obtained from 6 patients, selected externally; and 3) Biopsy samples from controls were obtained from different muscle groups than the eteplirsen-treated patients. For these reasons, we believe the control value of 0.08% dystrophin in untreated patients is uncertain, making the relative change in dystrophin difficult to estimate.

In any case, the level of dystrophin was 0.9% of normal after 3.5 years, such that, in absolute terms, the increase from baseline would be, at most, 0.9%, assuming a “worst case” for untreated patients, i.e., zero dystrophin. This finding is in sharp contrast to the value for percent dystrophin-positive fibers, which ranged from 14.2 to 20.0% for the eteplirsen-treated patients at Week 180. As discussed above, we believe that immunofluorescence analysis (percent positive fibers) is not a reliable method to quantify dystrophin content. Of note, the correlation between the two independent methods used to quantify dystrophin in muscle samples was weak (Figure 1).

Figure 1: Correlation between Two Methods Used to Quantify Dystrophin in Skeletal Muscle: Patients from Study 201/202



3) Clinical Effects Reflecting Muscle Function

Study 201/202 began as a 24-week randomized controlled study comparing three groups of patients treated weekly with intravenous eteplirsen 50 mg/kg, eteplirsen 30 mg/kg, or placebo (Study 201). After the randomized placebo-controlled phase, patients entered an open-label

extension phase, i.e., Study 202. Study 201 and Study 202, however, assessed the same patients, and *de facto* constitute two phases of the same study.

As noted above, the prospectively planned primary endpoint in Study 201 was the change from baseline in percent of dystrophin positive fibers in muscle tissue. The study had two pre-specified secondary endpoints: 1) change from baseline in CD3, CD4, and CD8 lymphocyte counts in muscle biopsy tissue at Week 12 or Week 24; and 2) change from baseline to Week 24 in 6-Minute Walk Test (6MWT).

For the prospectively planned analysis in Study 201, there was no statistically significant difference on the change from baseline to Week 24 in 6-Minute Walk (6MW) distance between eteplirsen 50 mg/kg, eteplirsen 30 mg/kg, and placebo.

Two patients in the 30 mg/kg group became unable to ambulate soon after the study start. The applicant then pooled the six remaining eteplirsen patients and compared them to the four placebo patients, an unplanned *post hoc* analysis. No nominally significant difference between eteplirsen and placebo was identified in that *post hoc* analysis.

The applicant conducted a number of additional *post hoc* analyses, comparing the 6 patients who received eteplirsen in the 24-week double-blind phase of Study 201 and could still ambulate at the end of Study 201 (and continued on open-label eteplirsen in Study 202) to those originally treated with placebo in the double-blind phase of Study 201, and later switched to open-label eteplirsen. Based on these analyses, the applicant stated² that “48 weeks of treatment with eteplirsen resulted in an unprecedented and clinically meaningful 67.3-meter clinical benefit on the 6MWT compared to placebo for 24 weeks followed by eteplirsen for 24 weeks.” Considering the *post hoc* nature of the analyses, the post-randomization exclusion of two patients who lost ambulation in Study 201, and the limitations of the open-label design for protecting against expectation bias on effort-dependent endpoints such as the 6MWT, FDA indicated to the applicant that data from Study 202, as presented, did not provide interpretable evidence of benefit.

FDA strongly encouraged the sponsor to conduct an adequately powered, randomized, placebo-controlled trial(s) to assess the clinical effect of eteplirsen. But in the context of an ongoing series of reports from the applicant and its academic associates describing marked effects on dystrophin production and stabilization of disease progression, many in the DMD community had strong reservations regarding the ethics and practicality of conducting another

² End-of-Phase 2 meeting of March 13, 2013.

placebo-controlled trial of eteplirsen. Given the apparent difficulty of doing such a trial, FDA expressed willingness to consider an externally controlled trial, although stating clearly that interpretation of the data could be difficult, and that the acceptability of the study would be a matter for NDA review.

The applicant proposed the submission of a New Drug Application relying on all available open-label data from Study 202 (up to Week 144) compared to a natural history cohort of untreated patients. FDA advised the applicant to identify external control groups appropriately matched to Study 202 patients, including similar treatment modalities, and to provide patient-level data. The applicant identified two DMD patient registries as a source of external data: the “Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry. In addition, FDA has very recently been able to obtain data from another patient registry, the Cooperative International Neuromuscular Research Group, and is in the process of conducting a separate analysis, the progress of which will be discussed at the advisory committee meeting. FDA is also working with investigators from the Centers for Disease Control and Prevention (CDC) and the Muscular Dystrophy Surveillance Tracking and Research Network (MD STARnet) to obtain additional data that may be informative about age of loss of ambulation in exon 51-skippable patients (6MW distance and other timed tests were not measured in this population-based tracking program).

The applicant conducted a *post hoc* comparison of the patients in Study 201/202 to data from the “Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry. They attempted to match patients in Study 202 with patients from these two registries based on five factors: 1) corticosteroid use at baseline (use/non-use); 2) sufficient longitudinal data for 6MWT available (Y/N); 3) age \geq 7 years (Y/N); 4) genotype amenable to any exon skipping therapy (Y/N); and 5) genotype amenable to exon 51 skipping therapy (Y/N). Patients did not have to match for baseline 6MW distance.

The problems of externally-controlled studies are well recognized. Under the proper circumstances, FDA regulations (21 CFR 314.126) recognize that historical control studies can be considered adequate and well-controlled studies, but there are many concerns with the interpretability of such studies. These are discussed in detail in international guidelines (International Conference on Harmonization Guideline, “Choice of Control Group and Related Issues in Clinical Trials” – ICH E10 [2000]).

FDA identified several issues related to the use of an external natural history that needed to be addressed prior to submission of the NDA.

First, FDA asked the applicant to establish that treatment modalities, including the physical therapy programs and steroid regimens used, were similar between patients from Study 201/202 and the externally-controlled population. There appear to be some important differences between the two groups. For example, the mean age of steroid start was over one year later in the control group than in eteplirsen-treated patients (age 6.4 years vs. 5.2 years). There were differences in steroid regimens used. The impact of these differences is impossible to estimate in the context of a non-randomized study.

Second, FDA noted that for most of its duration, Study 201/202 was open-label, with all patients receiving eteplirsen, and that performance on the 6-minute walk test could be influenced by expectation bias, motivation, and coaching. The patients in the external control group may not have been subject to these factors because they were not in a study and were not receiving an investigational therapy.

Third, although generally unavoidable, except in some cases where an externally controlled study is planned, external control groups are selected with data in hand. In this case, registries that served as the external control were identified and patient selection criteria were developed in February 2015, at a time when data on the 6-minute walk test were available in Study 201/202 for more than 3 years, and much of the data had already been generated in the external control group. A limited amount of the longitudinal data for the external control group was generated after selection of the patients, from February to December, 2015. The impact of these factors on the interpretability of the between-group comparisons cannot be determined.

Finally, although patients in external control groups can be matched based on factors that are known to be of prognostic importance, more concerning are factors that are unknown, yet potentially highly influential. Recent observational studies in DMD have been enrolling patients simultaneously with interventional trials of new drugs. Thus, patients in an observational cohort who were motivated to participate in an interventional drug study and who could qualify for enrollment might have dropped out of the observational study. With preferential loss of such subjects, patients who remained in the observational study may have been less motivated or less able to participate in interventional studies of new drugs, and in this sense, their prognosis could be worse. Although this reasoning is speculative, it highlights the difficulties in affirming that comparison of a group of patients in an interventional trial to an external group of control patients represents an “apples-to-apples” comparison. Comparisons can be confounded by unknown factors.

The applicant believes that the results of the external control comparison represent a result on an “intermediate clinical endpoint” – a clinical endpoint that can be measured earlier than irreversible morbidity or mortality (IMM), that is reasonably likely to predict an effect on IMM

or other clinical benefit, and that could suffice as a basis for accelerated approval. It should be noted that consideration for accelerated approval is based on the type of endpoints selected (surrogates; intermediate), and not on the adequacy of the studies supporting an effect on these endpoints. Thus, the evidence of an effect on an intermediate endpoint, if it is to serve as the basis for accelerated approval, must meet the evidentiary standard for substantial evidence from adequate and well-controlled studies. In this case, the externally-controlled study would need to be considered adequate and well-controlled to support full or accelerated approval.

The natural history in patients with DMD amenable to exon 51 skipping indicates a wide age range at the time of loss of ambulation, from 8 to 18 years of age for most patients. Progression in DMD occurs in a generally predictable stepwise fashion, with loss of ability to stand from the floor preceding loss of ability to walk independently, which precedes a decline in pulmonary function. Considering the entirety of the data submitted, we seek the Committee's opinion on whether there is convincing evidence that the clinical course of the 12 patients participating in Study 201/202 differs appreciably from the expected natural history of DMD, and, in light of the nature of the control group, whether a difference, if present, is interpretable.

The figures below illustrate the progression of functional deficits in eteplirsen-treated patients in Study 201/202. These figures include eteplirsen data up to Week 216 (Year 4), and for Figures 2, 3, 5, 6, and 7, historical control data from the "Italian DMD Registry" and the "Leuven Neuromuscular Reference Center" registry. The updated registry data were submitted by the applicant in January 2016, and were not available at the time of writing of the original briefing materials.

Comparing 6MW distance in eteplirsen-treated patients to control patients observed for a similar duration, the applicant describes nominally significant results in favor of eteplirsen in Study 202, with a difference of 148 meters compared to the external control at Year 3, and 162 meters at Year 4.

Figure 2 shows the change in 6MW distance over time, comparing eteplirsen-treated patients in Study 201/202 through Week 216 (Year 4), to historical controls from the "Italian DMD Registry" and the "Leuven Neuromuscular Reference Center" registry.

Figure 2: 6MWT vs. duration of observation in eteplirsen-treated patients in Study 201/202 and external control (“Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry)

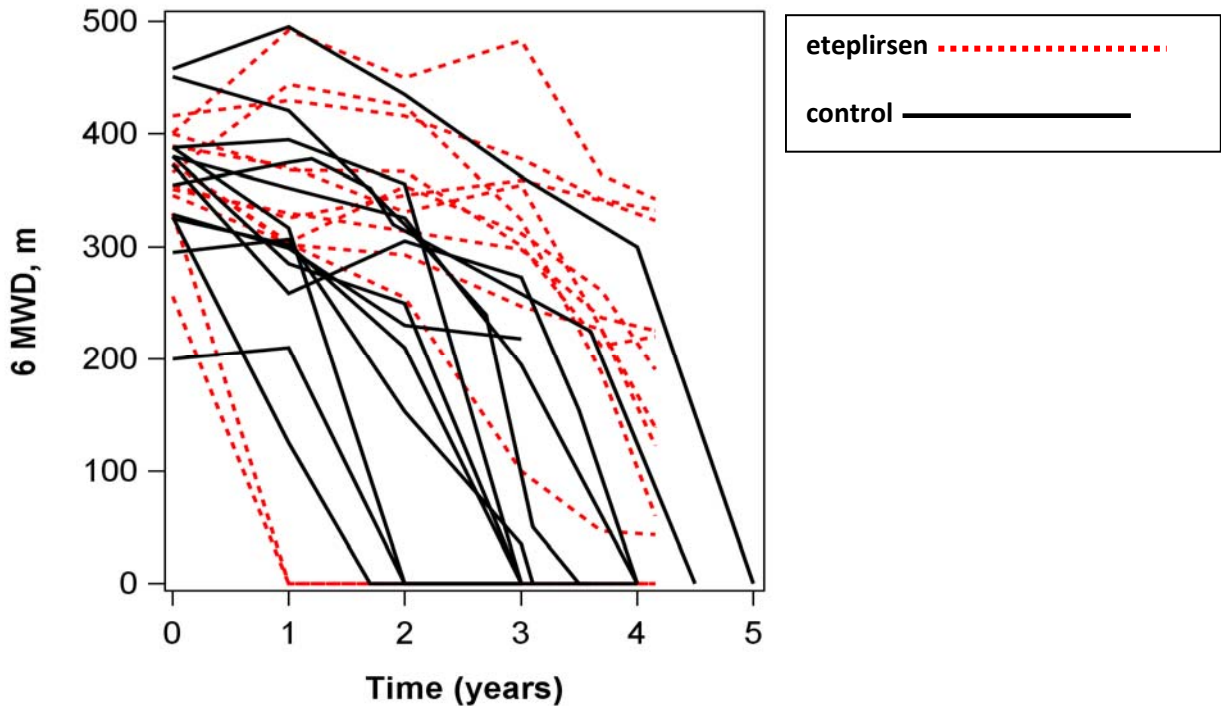
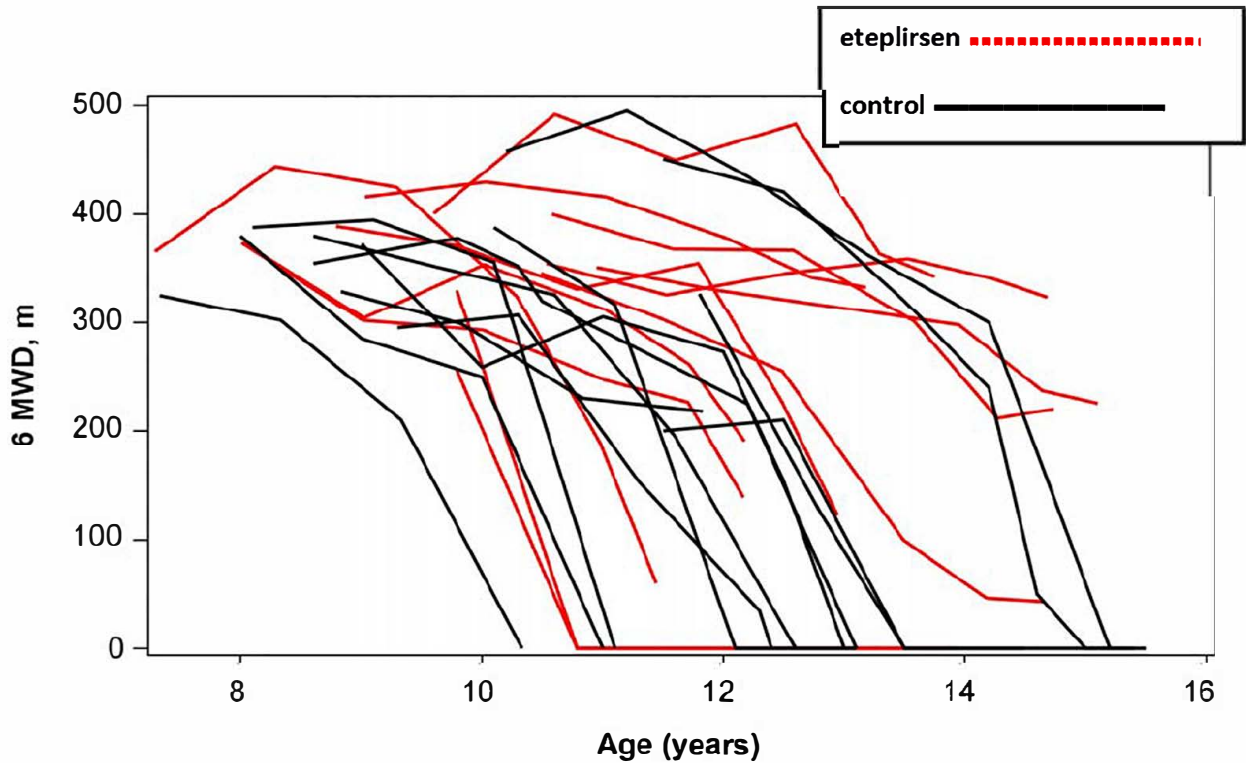


Figure 2 appears to show loss of ambulation in all patients in the historical control group by Year 5, whereas 10 of the 12 eteplirsen-treated patients are still ambulating at the Year 4 assessment. There are, however, several factors that raise questions as to whether there is a true difference in disease course between eteplirsen-treated patients and the control group:

1. Two patients in the historical control group who were reported to have lost ambulation nevertheless had 10-meter walk test values reported at the same points in time, providing evidence that ambulation was, in fact, not lost in these patients. This inconsistency highlights a concern that data collection and endpoint criteria may not have been applied similarly between eteplirsen-treated patients in Study 201/202 and the external control patients.
2. A comparison of 6MW distance versus age (as opposed to years on treatment), which takes into account the fact the functional abilities are correlated with age, shows substantial overlap between eteplirsen-treated patients and the historical control group (Figure 3).

Figure 3: 6MWT vs. age in eteplirsen-treated patients in Study 201/202 and external control (“Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry)



3. Eteplirsen-treated patients experienced the expected sequential worsening of functional abilities and muscle weakness, as demonstrated by the North Star Ambulatory Assessment (NSAA) scores. The NSAA is particularly important to the interpretation of the study results of Study 201/202. The NSAA has been specifically designed to measure functional ability in ambulatory patients with DMD, and can be used across a range of patient functional abilities. Among other functions, the NSAA measures activities of standing, walking, standing up from a chair, standing on one leg, climbing onto and descending from a box step, getting from lying to sitting, rising from the floor, jumping, hopping, and running. The NSAA is a comprehensive outcome measure, and arguably more fully reflects function in DMD than does the 6MWT. NSAA remains, however, dependent on subject effort, and is not immune to possible bias.

All eteplirsen-treated patients show progressive declines in NSAA scores, with 6 patients moving to NSAA scores that have been associated in the clinical setting with being one

year from stopping ambulation, and an additional four patients moving to scores associated with being within 2 years from stopping ambulation³ (Figure 4).

Figure 4: North Star Ambulatory Assessment (NSAA) scores vs. duration of observation in eteplirsen-treated patients in Study 201/202. The two horizontal lines indicate NSAA scores of 9 and 13, which have been reported to be associated with being either 1 or 2 years, respectively, from loss of ambulation.

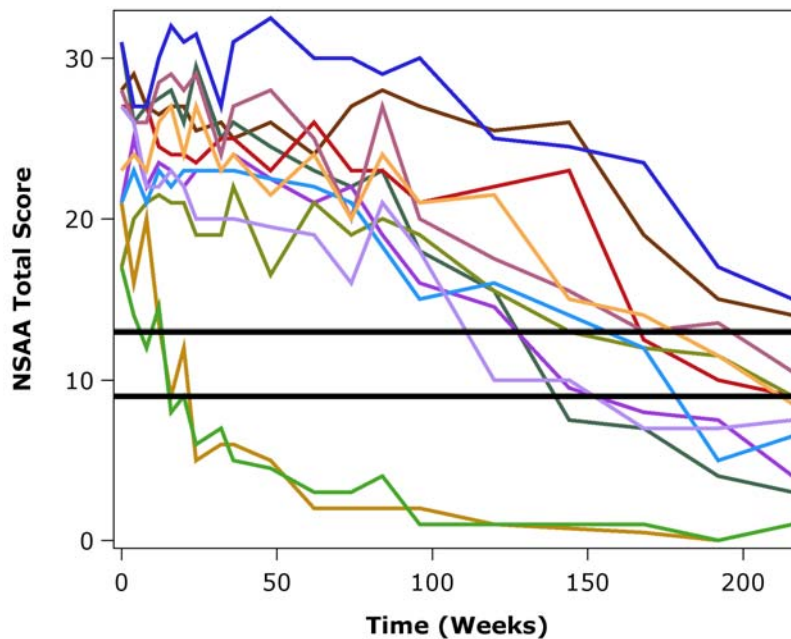


Figure 5 compares the NSAA data from eteplirsen-treated patients in Study 201/202 to those of the external control group. There is substantial overlap between eteplirsen-treated patients and the external controls, suggesting a similar disease course. The comparison of mean NSAA scores in Figure 5 (right) suggests that the external control patients started at lower baseline NSAA scores and have declined at the same rate as eteplirsen-treated patients. The overlap of the standard error bars is also notable.

A comparison of NSAA vs. age (Figure 6) also fails to show a substantial difference in disease course between eteplirsen-treated patients and the external controls.

³ Ricottii V, Ridout DA, Pane M, et al. The NorthStar Ambulatory Assessment in Duchenne muscular dystrophy: considerations for the design of clinical trials. *J Neurol Neurosurg Psychiatry* 2015,1-7

Figure 5: NSAA scores vs. duration of observation in eteplirsen-treated patients in Study 201/202 and external control (“Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry). Left: individual results. Right: mean results (\pm SD).

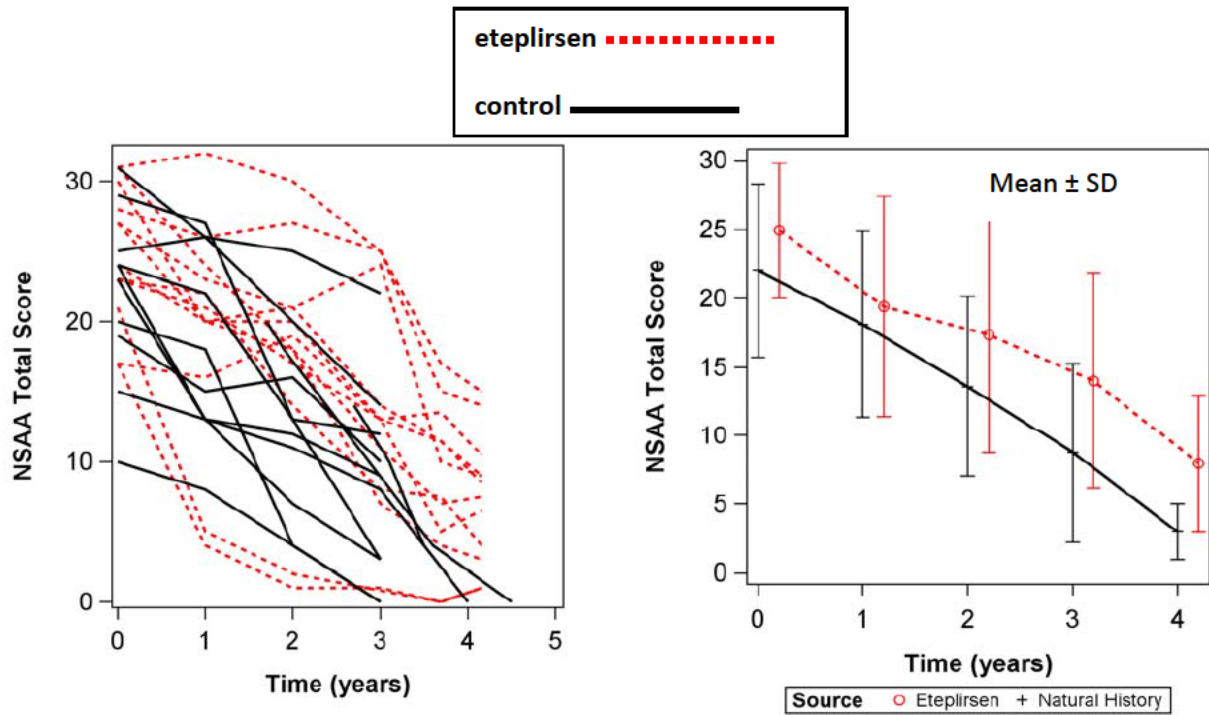
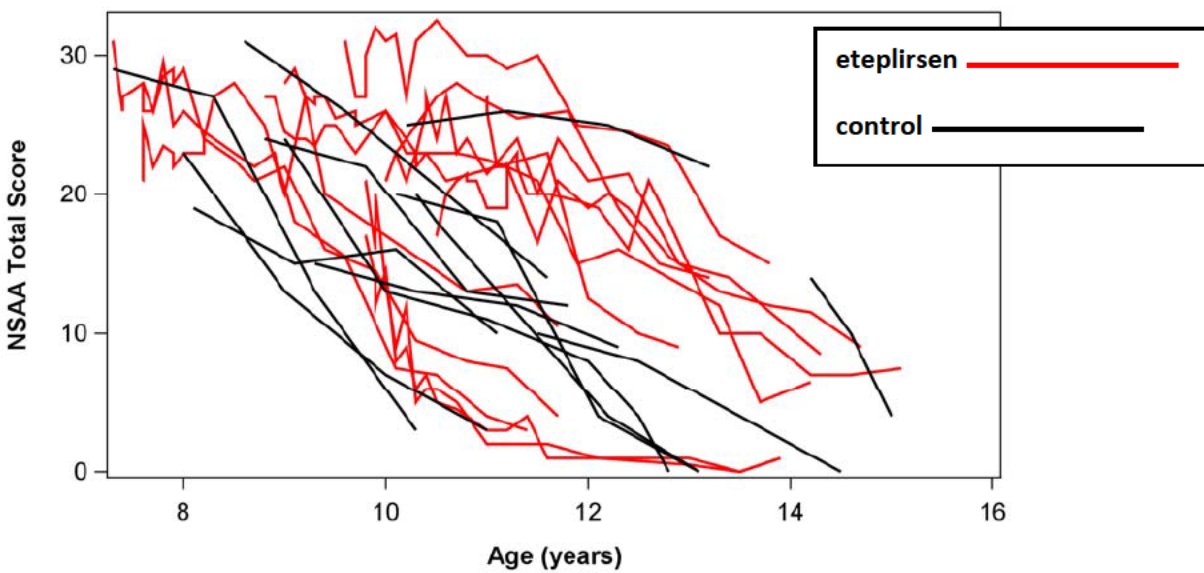
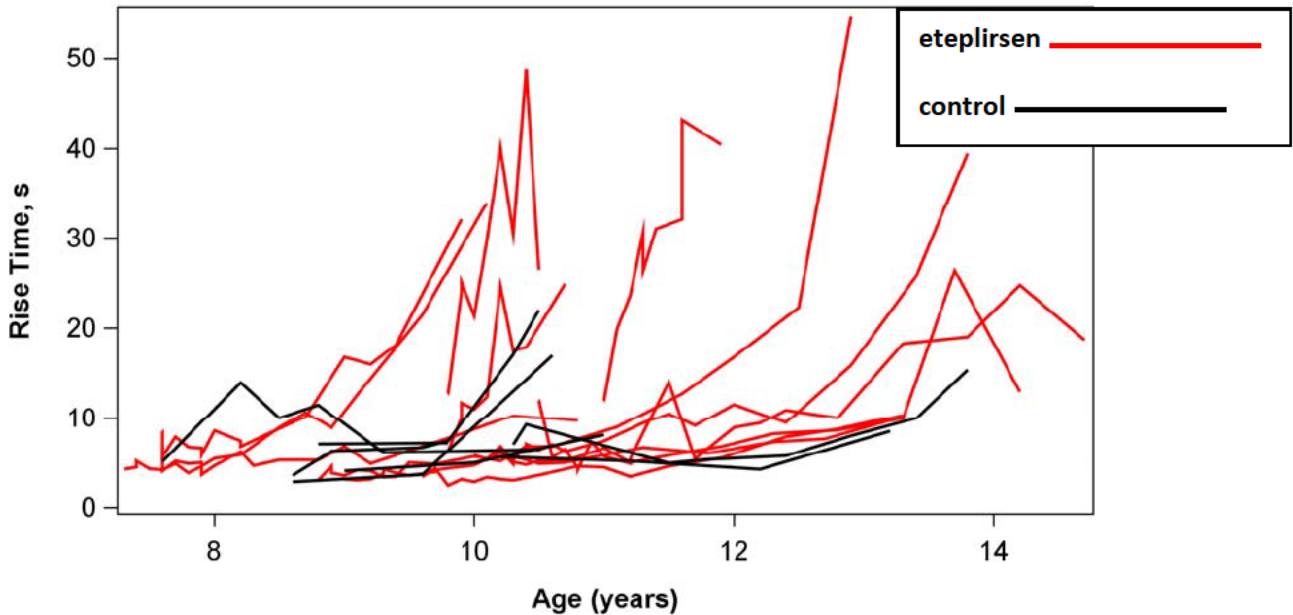


Figure 6: NSAA scores vs. age in eteplirsen-treated patients in Study 201/202 and external controls (“Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry)



- Most eteplirsen-treated patients had marked increases in rise time, and several became unable or nearly unable to rise from the floor, which predicts a high likelihood of loss of ambulation within 1 or 2 years,⁴ and illustrates substantial disease progression. In addition, progression of rise time was generally similar between eteplirsen and natural history cohort patients (Figure 7).

Figure 7: Rise time vs. age in eteplirsen-treated patients in Study 201/202 and external control (“Italian DMD Registry” and the “Leuven Neuromuscular Reference Center” registry)



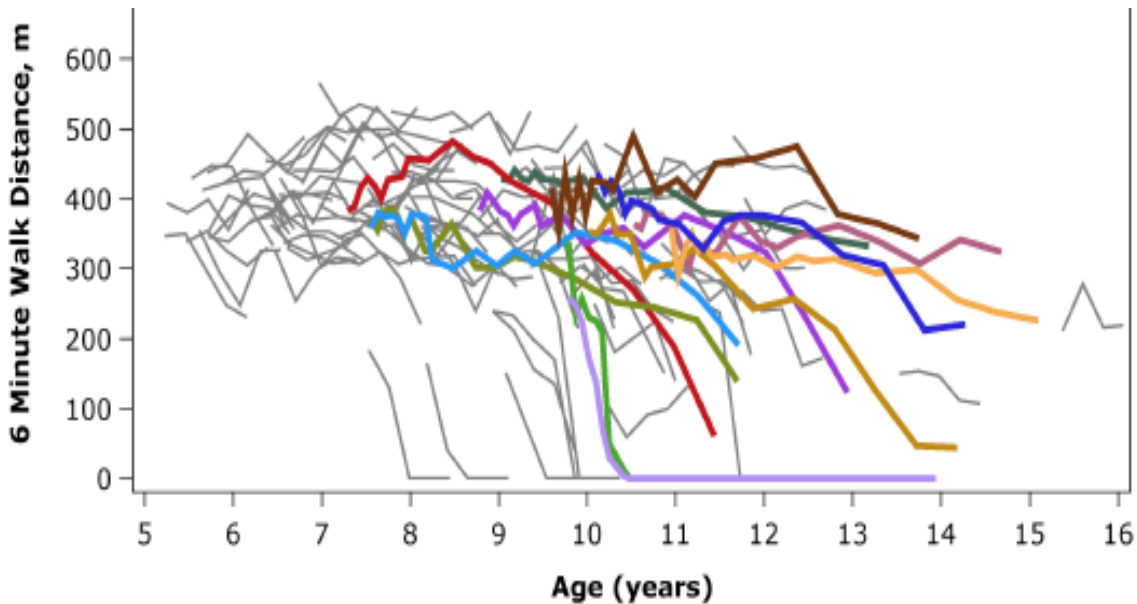
- The 6MW distances from eteplirsen-treated patients were compared to publicly available data from patients assigned to a placebo group in a recent large randomized controlled study in patients with DMD with mutations amenable to exon 51 skipping⁵ (Figure 8). Considering 6MW distance as a function of age, the values in eteplirsen-treated patients are within the range observed in the cohort of placebo patients from the other development program, again supporting a similar course between eteplirsen-treated patients and natural history. Although patients in the placebo-controlled group were each followed for only one year, these patients were in an age range similar to

⁴ McDonald CM et al (2013) The 6-minute walk test and other endpoints in Duchenne muscular dystrophy: longitudinal natural history observations over 48 weeks from a multicenter study. *Muscle & Nerve*. 48, 343-356

⁵<http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/UCM475956.pdf>

eteplirsen-treated patients, and provide a useful illustration of the range of 6MW distances that can be observed in DMD patients at various ages.

Figure 8: 6MWT in eteplirsen-treated patients in Study 201/202 (colored lines), compared to patients who received placebo in a recent randomized placebo controlled study in patients with mutations amenable to exon 51 skipping (grey lines)



Clinical Safety

To support marketing approval, the safety of a drug must be supported by an adequate number and duration of patient exposures to characterize the risks. Having said that, FDA will consider the serious and life-threatening nature of DMD and other severe dystrophinopathies when determining the minimum number and duration of patient exposures needed to assess safety.⁶ Drugs shown to provide an important benefit would generally need less safety data to provide adequate assurance that risks are commensurate with benefits.

No safety signal of significant concern has been identified for eteplirsen, although the clinical safety database for eteplirsen is small, as only 12 patients were exposed for one year or longer, and only 36 patients were exposed for 24 weeks or longer (the applicant included safety data

⁶ <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM450229.pdf>

from ongoing open-label studies). As a consequence, the one-year database only has adequate power to assess the frequencies of the more common adverse events (i.e., about 20% or greater). Less frequent events, possibly serious, may not have been identified to date because of the small database. Nevertheless, FDA recognizes that those affected by life-threatening and severely disabling illnesses with unmet medical need are generally willing to accept greater risks and greater uncertainty about risks.

Regulatory Requirements for Approval

Although approvability of a drug reflects a benefit-risk assessment, the decision about approvability is necessarily stepwise, requiring first that the drug is found to be effective, prior to consideration of benefit-risk.

The effectiveness requirement for a drug was added to the Food Drug and Cosmetic Act (FD&CA, the Act) in 1962. The 1962 amendments included a provision requiring manufacturers of drug products to establish a drug's effectiveness by "substantial evidence." Substantial evidence was defined in section 505(d) of the Act as:

"...evidence consisting of adequate and well-controlled investigations, by experts qualified by scientific training and experience to evaluate the effectiveness of the drug involved, on the basis of which it could be fairly and responsibly concluded by such experts that the drug will have the effect it purports or is represented to have under the conditions of use prescribed, recommended, or suggested in the labeling or proposed labeling thereof."

It has been FDA's position, based on the language of the statute and the legislative history of the 1962 amendments, that Congress generally intended to require at least two adequate and well-controlled trials, each convincing on its own, to establish effectiveness.

In 1997, under the FDA Modernization Act (FDAMA), section 505(d) of the Act was amended to make it clear that the Agency may consider "data from one adequate and well-controlled clinical investigation and confirmatory evidence" to constitute substantial evidence if FDA determines that such data and evidence are sufficient to establish effectiveness.

Thus, a single highly persuasive positive trial combined with independent findings that substantiate efficacy might support approval, but it is critical that the possibility of an incorrect outcome be considered and that all the available data be examined for their potential to either

support or undercut reliance on a single trial. FDA described in a guidance document⁷ the characteristics of a single adequate and well-controlled study that could support an effectiveness claim. These include: 1) large multicenter study; 2) consistency across study subsets; 3) multiple studies within a study (e.g., properly designed factorial study analyzed as a series of pairwise comparisons); 4) multiple endpoints involving different events; and 5) statistically very persuasive findings. Some of these characteristics largely pertain to more common diseases. DMD is a rare and serious disease without approved treatments, and FDA has long stressed that it is appropriate to exercise the broadest flexibility in applying the statutory standards to drugs for such diseases, while preserving appropriate guarantees for effectiveness and safety.⁸

Accelerated Approval

The applicant is seeking accelerated approval for eteplirsen. Accelerated approval is a particular type of approval that FDA may grant for a product for a serious or life-threatening disease or condition upon a determination that the product has an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit, or on a clinical endpoint that can be measured earlier than irreversible morbidity or mortality and that is reasonably likely to predict an effect on irreversible morbidity or mortality or other clinical benefit, taking into account the severity, rarity, or prevalence of the condition and the availability or lack of alternative treatments.⁹

Two potential pathways to accelerated approval were discussed with the applicant during the eteplirsen development program:

1. Using clinical data from Study 201/202 on 6-minute walk distance as an intermediate clinical endpoint that could have the potential to support accelerated approval.

Under that approach, the basis for accelerated approval would be a conclusion that eteplirsen reduced the rate of decline of walking performance to an extent that is reasonably likely to predict a long-term beneficial effect on irreversible morbidity or mortality. It should be noted,

⁷ US Food and Drug Administration. Guidance for Industry. Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products. <http://www.fda.gov/downloads/Drugs/.../Guidances/ucm078749.pdf>. May 1998. Accessed December 17, 2015.

⁸ 21 CFR 312.80, subpart E

⁹ <http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm358301.pdf>

however, that FDA would consider an effect on walking distance to be a clinical benefit that, if demonstrated, would support full approval. Study 201 failed to show an advantage of eteplirsen over placebo on 6-minute walk distance during the placebo-controlled phase of the trial. The specific finding proposed by the applicant as supporting accelerated approval is the comparison of 6-minute walk distance between the 12 patients in Study 201/202 and external controls, where the control patients were selected *post hoc*. Again, as externally controlled trials can be considered well-controlled studies (21 CFR 314.126), if the data were considered reliable they would support full, not accelerated, approval. There are significant concerns regarding the ability to draw valid conclusions from this externally controlled comparison. Moreover, comparisons between patients in Study 201/202 and patients in a related development program who had received placebo suggest that the change in 6-minute walk distance with eteplirsen was consistent with the natural history of the disease.

2. Using dystrophin data as a surrogate endpoint to support accelerated approval.

FDA indicated in the draft DMD guidance⁸ that biomarkers that reliably reflect the health and amount of skeletal muscle may, if supported by sufficient scientific evidence and acceptable analytical methods, be used as surrogate endpoints to support accelerated approval of a new DMD drug. Such a biomarker would have to be “reasonably likely to predict clinical benefit” in order to be acceptable as a basis for accelerated approval.

For eteplirsen, the quantification of dystrophin present in the fourth muscle biopsy was assessed by Western Blot, and compared with treatment-naïve controls that were selected by the applicant. The dystrophin level in patients who had been treated with eteplirsen for some 3.5 years was 0.93% ± 0.84% (mean ± standard deviation) of normal, far below levels generally observed in a milder form of muscular dystrophy known as Becker-type muscular dystrophy (BMD). The minimum level of dystrophin that might be reasonably likely to predict clinical benefit in patients with BMD remains unknown, but experts in DMD^{10,11} have stated that levels less than 3% of that of normal healthy muscle are generally associated with the typical DMD phenotype, and have proposed that “induction of approximately 10% of normal dystrophin levels sets a minimum level to confer measurable clinical benefit.”¹²

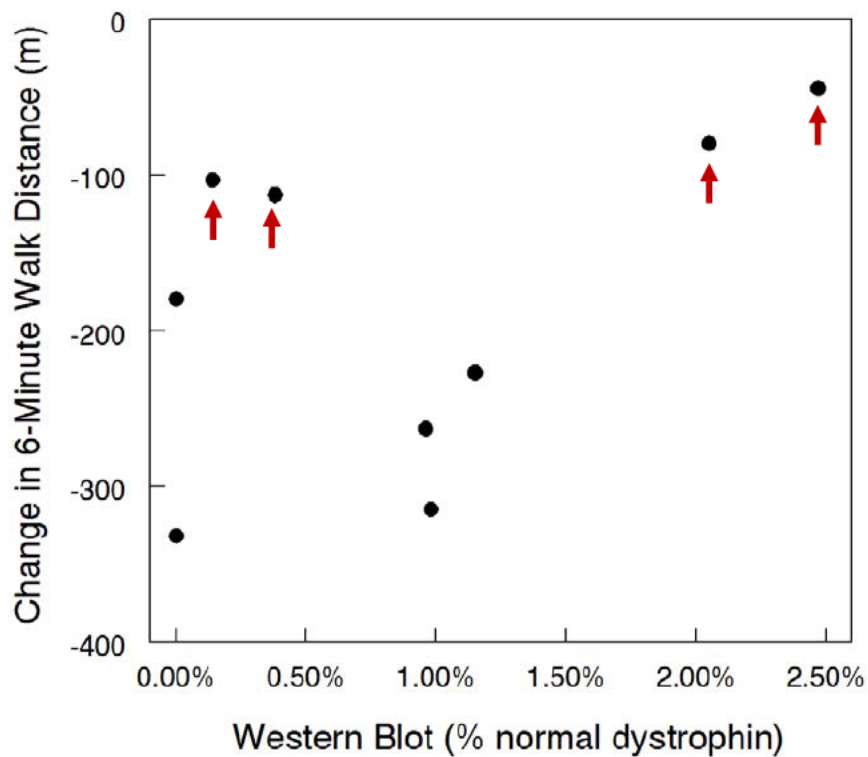
¹⁰ Flanigan KM (2014) Duchenne and Becker muscular dystrophies. *Neurol Clin.* 32,1 671-688

¹¹ Lu QL, Cirak S, Partridge T (2014) What can we learn from clinical trials of exon skipping for DMD? *Mol Ther Nucleic acids.* 3, e152

¹² Wilton SD, Veedu RN, Fletcher S (2015) The emperor’s new dystrophin: finding sense in the noise. *Trends in Molecular Medicine.* 21, 417-426

If production of dystrophin protein is “reasonably likely to predict clinical benefit,” then one would expect a correlation between the level of dystrophin and ambulation (as assessed by the 6MWT) in eteplirsen-treated patients. In Study 201/202, there are too few patients to perform a rigorous analysis. But for the 9 patients who were able to ambulate and who volunteered to undergo a biopsy at Week 180, it is apparent that for the 4 patients whose 6MW distances were best preserved (red arrows, Figure 9), 2 had very low levels of dystrophin, and 2 had the highest levels. Thus, there is no apparent correlation between 6MWT and dystrophin levels in eteplirsen-treated patients.

Figure 9: Study 201/202; change in 6-minute walk distance (Week 180 minus Baseline) versus dystrophin level as determined by Western blot. (Two patients who lost ambulation are omitted.)



Importantly, the evidentiary standards for effectiveness are not lower for biomarker endpoints used to support accelerated approval, nor should accelerated approval be used to compensate for weak or inconsistent clinical findings.

Although FDA is prepared to be flexible with respect to a devastating illness with no treatment options, flexibility does not mean approving drugs for which substantial evidence of effectiveness has not been established. Thus, as you digest the background materials, we hope you will carefully consider the strengths and weaknesses of all of the data, and be prepared to consider and discuss whether or not you believe that efficacy has been established.

It is important to recognize that no final conclusions have been reached on the approvability of this application, and we look forward to a fruitful discussion of these issues at the Advisory Committee Meeting on April 25, 2016.

II. Drafts Points To Consider

FOOD AND DRUG ADMINISTRATION (FDA)
Center for Drug Evaluation and Research (CDER)

Peripheral and Central Nervous System Drugs Advisory Committee Meeting

DRAFT POINTS TO CONSIDER

April 25, 2016

1. Consider the data for dystrophin expression, including the following
 - a. Experimental methods, including consideration of accuracy, reliability, reproducibility, etc.
 - b. Potential clinical meaning, including consideration of amount of dystrophin relative to patients with Becker muscular dystrophy, functionality of the truncated dystrophin, and percent of muscle fibers with detectable dystrophin.
2. Consider the data for clinical measures, including the following
 - a. Design and potential interpretability of Study 201/202, including consideration of a) the placebo-controlled period, and b) comparison of the open-label experience to natural history.
 - b. Results of Study 201/202 in the context of the study design.
3. Consider the possible design of any future efficacy and safety studies that might be necessary.

III. Clinical Team Leader Memorandum to the Committee

MEMORANDUM

DATE: March 29, 2016

FROM: Ronald Farkas, M.D., Ph.D.
Clinical Team Leader
Division of Neurology Products, CDER, FDA

TO: Members and Invited Guests of the Peripheral and Central Nervous Systems
Drugs Advisory Committee (PCNS AC)

SUBJECT: Clinical Team Leader Memorandum for New Drug Application (NDA) 206488, for
the use of Exondys 51 (eteplirsen) for the treatment of Duchenne muscular
dystrophy in patients with mutations amenable to exon 51 skipping

Preface

This memorandum is revised from the memorandum for the PCNS AC meeting for eteplirsen that had been scheduled for January 22, 2016. The revisions are based on additional data submitted by the applicant for both eteplirsen-treated and natural history patients, newly available natural history from the Cooperative International Neuromuscular Research Group (CINRG), new analyses of data previously submitted by the applicant, and comments from other interested parties subsequent to the release of the previous memorandum. Following release of the FDA briefing material the applicant stated in an addendum¹ that there were key inaccuracies in the FDA material regarding dystrophin analytical methodology and findings. FDA's responses to the applicant's statements are also included in this document (the applicant's table of "Key Inaccuracies" is appended to this document). For clarity, this memorandum contains the previous text and figures, with new text in italics.

¹<http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/UCM481913.pdf>

Contents

1. Disease Background	4
2. Eteplirsen Drug Development Rationale.....	4
3. Dystrophin Evidence	5
a. Study 33	5
b. Study 28	5
c. Study 201/202, First 3 Biopsies.....	6
Immunofluorescence images (Study 201/202, first 3 biopsies)	7
Western blots (Study 201/202, first 3 biopsies)	11
d. Study 201/202, 4th Biopsy.....	11
Exon Skipping	13
Western Blot, 4 th biopsy.....	14
Percent Positive Fibers.....	15
Total Dystrophin Immunofluorescence Intensity	16
e. Dystrophin in BMD.....	17
Quantity	17
Timing.....	18
Functionality	18
Localization	19
Unusual BMD Patients	19
f. Reviewer Discussion, Dystrophin Quantification Methods	20
g. FDA Review Team Preliminary Conclusions on Dystrophin Findings.....	22
Clinical Efficacy Evidence	23
a. Design and analysis of Study 201/202	23

b.	Rate of progression of 6MWT in eteplirsen-treated patients is consistent with expected natural history	31
c.	Increases in rise time in eteplirsen-treated patients predict a high likelihood of sequential loss of ambulation within 1 or 2 years	42
d.	Issues with comparison of eteplirsen-treated patients with applicant’s proposed historical controls	45
e.	NSAA, Eteplirsen vs. Applicant’s Controls.....	50
f.	6MWT, Eteplirsen vs. Applicant’s Controls.....	52
g.	FDA Review Team Preliminary Conclusions, Clinical Endpoints	59
4.	Clinical Safety	62
	Appendix: Applicant’s table of “Key Inaccuracies in the FDA Briefing Document”	63

1. Disease Background

DMD is caused by genetic mutations in the dystrophin gene that result in near absence of the dystrophin protein from muscle. Dystrophin is thought to maintain the structural integrity of the muscle cell membrane by connecting the cytoskeleton to the surrounding extracellular matrix, and to act as a scaffold for several signaling molecules that also contribute to normal muscle physiology. Immunological and inflammatory processes downstream of dystrophin deficiency appear to contribute to muscle pathology in DMD. Key manifestations of DMD include progressive degeneration of skeletal and cardiac muscle resulting in loss of function in childhood and adolescence and premature death from respiratory or cardiac failure in the second to fourth decade. Corticosteroid therapy is considered standard of care, delaying loss of ambulation and respiratory decline by several years.

2. Eteplirsen Drug Development Rationale

Because of the near total lack of dystrophin in DMD, one rational approach to therapy involves trying to restore dystrophin expression. In many patients with DMD, very small amounts of a shorter than normal “truncated” form of dystrophin are produced, due to what might otherwise be considered an error in mRNA splicing: an exon is left out, or “skipped”, which, in the setting of specific DMD-causing mutations, can result in restoration of the mRNA reading frame. Unfortunately, the small amount of exon skipping that occurs naturally in DMD patients does not appear to appreciably slow muscle degeneration. It was reasoned, however, that if exon skipping could be augmented by drug therapy, levels of the truncated dystrophin could be increased to a level high enough to confer clinical benefit. Eteplirsen was designed to bind to dystrophin mRNA at a specific site to cause the splicing machinery to skip exon 51, thus restoring the dystrophin reading frame in certain amenable patients, and increasing production of the truncated dystrophin. How *much* of the truncated dystrophin would be necessary to confer clinical benefit remains an open question, but a related form of muscular dystrophy, called Becker muscular dystrophy (BMD), provides a natural model of what exon skipping in DMD might achieve. In so-called “exon 51-model” BMD patients, the same truncated form of dystrophin that would be produced by eteplirsen in DMD patients occurs naturally. These BMD patients experience a mild, or in some cases asymptomatic, muscle disease. Importantly,

however, the truncated dystrophin in these BMD patients is expressed at high levels, roughly 50- to 100%² of what would be expected for normal dystrophin.

3. Dystrophin Evidence

Dr. Ashutosh Rao, from the Office of Biotechnology Products, reviewed dystrophin methodologies and supporting assays. The effect of eteplirsen on dystrophin expression was examined in 3 clinical studies: Study 33, Study 28, and Study 201/202, as follows:

- a. Study 33:** In this exploratory phase 1 study, small doses of eteplirsen (up to 0.9 mg total) were injected directly into a foot muscle in 7 patients with DMD. An increase in dystrophin expression was reported adjacent to the needle track, but it is not clear whether, or to what degree, this might reflect the activity of eteplirsen when given by the intravenous (IV) route, which does not produce similar high local concentrations or mechanical effects.
- b. Study 28:** In this exploratory study, eteplirsen was administered intravenously once a week for 12 weeks at doses ranging from 0.5 to 20 mg/kg, with up to 4 patients per dose level. The methods for dystrophin quantification were not reviewed by FDA prior to the conduct of the study, and FDA has concerns about the reliability of the methods and procedures. In one response from the applicant to an information request from FDA about quality control methods, the applicant responded that “Study 28 was an exploratory phase 1b study which was only intended to generate proof of concept data to guide future studies. For this reason, quality controls for the dystrophin data in Study 28 were not properly optimized.” In addition, Study 28 examined dystrophin levels after 12 weeks of dosing, but it is necessary to understand dystrophin levels that are present with longer-term, more clinically relevant durations of therapy. Thus, as described below, FDA considers the 4th biopsy from patients in Study 201/202, which was taken after 180-weeks of treatment with eteplirsen, to be of greater potential clinical relevance.

The results of Study 28 do not appear to be interpretable. Western blot bands were too saturated to allow reliable quantification. Study design and conduct issues were also a major concern. The study was unblinded and, according to the applicant, assays were repeated and reanalyzed. Repeating assays and analyses when unblinded to treatment can increase the risk of bias and false positive findings; results supportive of the preferred hypothesis may be preferentially selected, whereas ambiguous or non-

² Anthony K et al (2011) Dystrophin quantification and clinical correlations in Becker muscular dystrophy: implications for clinical trials. Brain. 134,3544-3556.

supportive results may be discounted as having resulted from the types of technical failures that are common in laboratory research. The Study 28 report from the applicant states the following regarding repeated assays and analyses: “Of note, the laboratory performing the Western blot analyses used multiple samples from the same patients to re-analyze the results. Initially, the Western blot analyses reported the results from one sample per patient and any post-treatment increases in dystrophin protein level were reported as an ‘X’-fold increase from baseline. Subsequently, while preparing the Lancet publication, the laboratory repeated several Western blots to achieve publication standard results and also to test different pieces of muscle within a patient. These results were reported as the maximum amount of dystrophin per patient and were expressed as a percentage of normal.”

As detailed in later sections of this memo, dystrophin levels in the 4th biopsies of Study 201/202, which were obtained after 180 weeks of eteplirsen treatment, were estimated to be about 0.9% of the amount in normal muscle. In contrast, Study 28 reported amounts 10- to 20-fold higher after only 12 weeks of eteplirsen treatment, in patients treated with doses of eteplirsen as low as 1/10th those used in study 201/202. In light of the issues noted above, however, FDA does not believe the dystrophin results from Study 28 are interpretable

- c. Study 201/202, First 3 Biopsies:** Study 201/202 was a 3-arm, 12-patient study comparing the effects of 30 mg/kg or 50 mg/kg IV eteplirsen to placebo. Biopsies were taken at baseline, week 12 (for half the patients), week 24 (for the other half), and week 48 for all patients. During the development of eteplirsen FDA communicated to the applicant concerns about the biomarker studies on the first 3 biopsies.³ With additional review following submission of the NDA, it is not clear that any of the dystrophin biomarker data from the first 3 biopsies are reliable or interpretable.

³ e.g. at a meeting on March 13, 2013, FDA stated “while we do not believe that you have adequately characterized the quantity of truncated dystrophin produced by eteplirsen treatment (Western blot data is not available), the immunofluorescence data you presented suggest that a much lower quantity of truncated dystrophin is produced by eteplirsen treatment than is present in BMD.” In the April 15, 2014, advice letter in which potential pathways for approval were discussed, FDA stated “After examining the source data and images you provided in support of dystrophin protein expression from eteplirsen treatment, we remain skeptical about the persuasiveness of the data, and concerned about serious methodological problems explained previously.”

Immunofluorescence images (Study 201/202, first 3 biopsies)

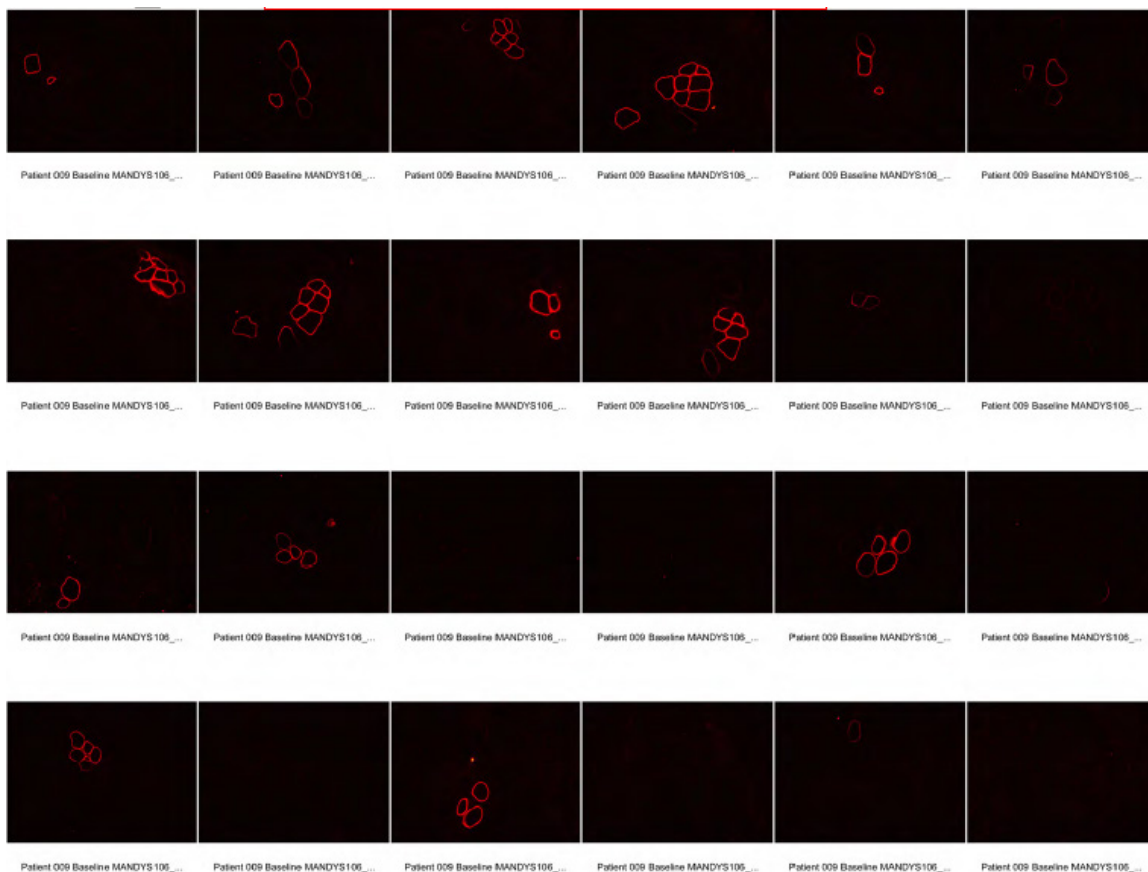
The measurement of total dystrophin immunofluorescence by Bioquant was first carried out on blinded baseline, Week 12, and Week 24 images, captured at 20x magnification. The results showed essentially no change in intensity for any patient. Negative results were obtained both when the study was conducted with MANDYS106 antibody or with Dys2 antibody. However, investigators attributed the negative results to the image magnification, and captured new images at 40x magnification after the blind was broken, with personnel reporting to FDA site inspectors that positive fields were uniquely selected for further quantitation. The images selected at 40x magnification showed roughly a doubling of immunofluorescence intensity for all patients between baseline and Week 12 (50 mg/kg patients) or week 24 (30 mg/kg patients). Because the analyses were intentionally targeted to fibers whose staining intensity exceeded a particular threshold, it is not clear whether these results are representative or interpretable.

The 20x immunofluorescence images on samples obtained through Week 24 were selected by an individual blinded to treatment group, but the microscopic fields to be photographed were selected manually by the operator, as opposed to a more automated method introduced for studies of the 4th biopsy. Bias in field selection may have resulted in preferential capture of bright fibers that appear similar to revertant fibers. Figure 1 shows all 24 fields captured from a single patient at Week 24 in Study 201. Three of the fields show a cluster of what appear to be the same revertant fibers that appear to extend through multiple levels of the tissue sample. Similar apparent over-representation of bundles of likely revertant fibers occurred for many other patients and time points; for example, images obtained at baseline from a different patient are shown in Figure 2.

Figure 1: Example of immunofluorescence fields, Study 201



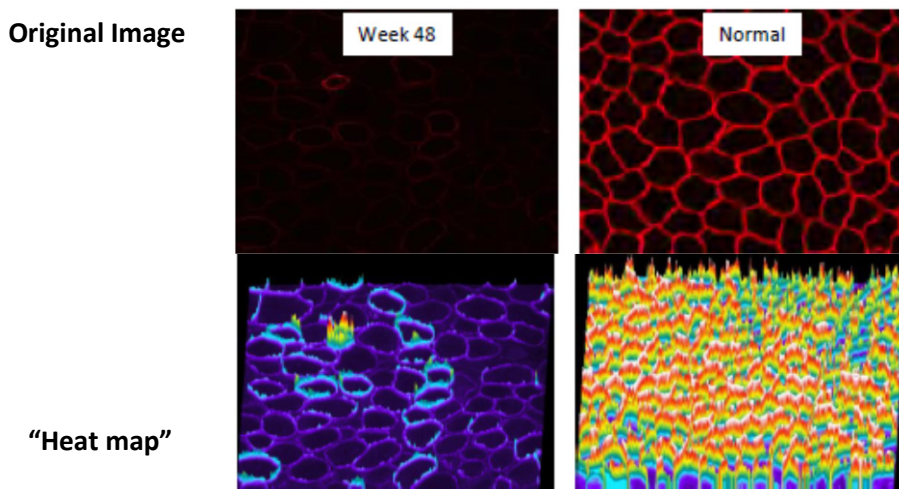
Figure 2: Example of Baseline Dystrophin Immunofluorescence



Week 48 samples were processed separately for dystrophin immunofluorescence from earlier samples, and had higher background staining. As a consequence, valid comparison is not possible with earlier time points for percent positive fibers or total immunofluorescence because the higher background staining, and not necessarily an effect of drug, could be responsible for any differences observed.

Importantly, the Week 48 immunofluorescence was still very low, and much less intense than normal controls, as shown in Figure 3. The top two images show the intensity as originally captured, and the bottom two images show the intensity converted to “heatmap” images that represent the observed (unmodified) pixel intensity as color, from low intensity blue to high intensity red and white.

Figure 3: Dystrophin Immunofluorescence vs. Normal Control



It is important to note that the applicant digitally processed⁴ dystrophin images in their background material (images in Appendix 12) in such a way that low intensity values were preferentially increased to produce a higher intensity and higher contrast image.

Note: following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the above paragraph as a key inaccuracy:

Sarepta: “The digitally processed images referenced by FDA in this statement were included in Sarepta’s briefing document for demonstration purposes only, and it is far more important to note that the referenced images were not used in the analysis of fiber intensity, nor to score dystrophin-positive fibers.”

FDA response: FDA acknowledges that digitally manipulated images were not used in the applicant’s numerical assessment of fiber intensity or percent positive fibers, but it is concerning that images used to provide evidence of an effect of eteplirsen greatly exaggerate the immunofluorescence signal from the muscle samples.

⁴ Per the applicant: To generate the enhanced inverted_b base100 Image (InvertBase100), the algorithm produces a non-linear mapping of r,g,b fluorescent values that will specifically enhance low contrast objects in the image. It does this by scaling the r,g,b fluorescent values using the following formula: $I' = 1 - 100^{(-I)}$ normalized by the max value of $1 - 100^{(-1)}$ for each of the channels independently. This results in low intensity values being stretched and therefore perceived as having a higher intensity and a higher contrast

Western blots (Study 201/202, first 3 biopsies)

Western blots from the first 3 biopsies are not considered interpretable because of substantial technical shortcomings, including lack of a dilution-series of normal muscle as a comparative control, saturation of bands such that ratios of intensity are unreliable and, in many blots, multiple bands in the region of dystrophin immunoreactivity that decrease confidence that the correct band was identified for quantification. Additional potential for bias was introduced because multiple Western blots were performed, with a number of different antibodies (Mandys106, Dys1, Dys2), with negative findings on many blots attributed to technical issues, whereas positive findings were attributed to drug effect.

d. Study 201/202, 4th Biopsy

Biomarker studies on the 4th biopsy obtained at Week 180 were conducted by the applicant with technical advice from FDA. However, the reliability of results remains questionable for a number of reasons, including the following:

- **Controls were not matched by muscle group:** Biopsies at Week 180 were taken from deltoid, one of the few muscle groups that, along with the calf muscle, can be hypertrophied in DMD.⁵ In contrast, both the baseline samples available from eteplirsen-treated patients, and most of the new external controls from untreated patients, were obtained from biceps (except for one, which was obtained from deltoid). There is little human data on differences in dystrophin levels between muscle groups in DMD but, in nonclinical models of DMD, there is evidence that dystrophin levels vary between muscles,⁶ which may affect the readout of experiments in which the effectiveness of the treatment is not particularly high.
- **Controls were not matched by patient:** There appears to be considerable inter-patient variability in dystrophin levels present in exon-51 skippable DMD. In Western blots from biopsies of extensor digitorum brevis (EDB),⁷ dystrophin levels averaged about 0.3% of normal, but ranged from undetectable to \approx 1% of normal or somewhat higher. The applicant obtained data from biopsies of 9 untreated patients, and reported an average

⁵ Pradhan S (2002) Valley sign in Duchenne muscular dystrophy: importance in patients with inconspicuous calves. *Neurol India*. 50,184-186.

⁶ Pigozzo S et al (2013) Revertant fibers in the mdx murine model of Duchenne muscular dystrophy: an age- and muscle-related reappraisal. *PLOS ONE*. 8,e72147.

⁷ FDA Advisory Committee presentation for drisapersen, slide 43.

dystrophin level of 0.08%.⁸ However, such a small sample size may not provide a reliable estimate of baseline levels that were present in the eteplirsen-treated patients. The dystrophin level estimated in these biceps controls is lower than the estimate from the EDB biopsies, perhaps because dystrophin levels truly differ between these muscle groups, or perhaps only secondary to chance when a small number of observations with high variability are compared.

- **Lack of independent confirmation:** The applicant has not obtained independent confirmation of dystrophin findings.⁹

***Note:** Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the above information as a key inaccuracy.*

***Sarepta:** “Methodology for dystrophin analyses of the fourth biopsy tissue samples, including confirmatory assessments of percent dystrophin-positive fibers (PDPF) analysis performed by 3 independent pathologists, were agreed with FDA prior to conducting any analyses of the fourth biopsy tissue samples. In accordance with the mutually agreed-upon protocols for the assessment of dystrophin-positive fibers in DMD muscle biopsy samples from the fourth biopsy obtained at Week 180, 3 independent pathologists performed a blinded assessment of the randomized muscle fiber microscopy images, which independently confirmed the results obtained by the pathologist at Nationwide Children’s Hospital (NCH).*

Assessment of PDPF at NCH indicated a significant increase in PDPF score ($p < 0.001$) relative to untreated control samples. This increase in PDPF score was confirmed by the 3 independent pathologists ($p < 0.001$).”

⁸ Noting, however, that values $< 0.25\%$ were rounded to zero. Including those lower values leads to an average level about twice as high, but still half as much as in EDB.

⁹ For example, in the April 15, 2014, letter discussing data that would be filed with the NDA, FDA stated “We expect that the initial biomarker data from these [newly exposed patients] exposures will start becoming available at about the time of NDA submission and shortly thereafter.” Also, as early as the July 23, 2013 meeting FDA expressed concern that “all muscle biopsies were obtained and processed by a single technician at a single study center” and that in part because of concern about bias, “we also ask that you confirm, [biomarker results] by an independent laboratory.”

FDA Response: *The FDA statement that biomarker studies on the 4th biopsy are considered of questionable reliability is correct. FDA explained to the applicant that it would be reasonable for them to perform the proposed analyses on the newly acquired biopsy tissue but that there were shortcomings and limitations to potential interpretability (communicated March 30, 2015):*

- **Controls for 4th biopsy:** *Prior to conduct of biomarker studies on the 4th biopsy, FDA provided the following advice about the shortcomings of the controls selected by the applicant and limitations the controls would place on interpretability:*
 - *“The control biopsy tissue that you propose to use is from a number of different muscle groups, such that differences that may exist in dystrophin expression among muscle groups may affect your results. However, in the context of other major sources of variability among biopsies (including both intra- and interindividual differences even within the same muscle group), it appears reasonable for you to proceed with these controls, with the understanding that dystrophin changes would need to be robust to be interpretable as a drug effect.”*
- **Meaning of Percent Dystrophin Positive Fibers (PDPF):** *FDA also reminded the applicant at that time of the importance of WB data for quantifying dystrophin:*
 - *“As proposed, your western blot method is likely to be more reliable for quantitative measurement of dystrophin.”*

Meaning of independent confirmation of findings: *Multiple readings of data from a single study, e.g., 3 independent readings of dystrophin-positive fibers, do not constitute an independent study. As early as the July 23, 2013 meeting FDA expressed concern with the applicant that “all muscle biopsies were obtained and processed by a single technician at a single study center.”*

Exon Skipping

The applicant reported positive findings for all patients on detection of exon 51-skipped mRNA, as measured by RT-PCR. However, RT-PCR is highly sensitive to the presence of even a few molecules of mRNA, and does not indicate how much, or even whether, any dystrophin protein might have been produced.

Western Blot, 4th biopsy

Western blot results for eteplirsen-treated patients are shown in Table 1. Dystrophin levels in treated patients were, on average, about 0.9% of normal¹⁰ (range <0.25% -2.5%) as measured by Western blot, the most quantitative method used by the applicant.

At the low dystrophin levels present in the Week 180 biopsies, random measurement error can be large in comparison to the estimated amount of dystrophin. Consequently, little confidence can be placed on any individual patient value, and the data should not be considered as reliable evidence that some patients failed to produce any dystrophin from eteplirsen whereas others were more responsive.

Note: *Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statement, “Random measurement error can be large in comparison to the estimated amount of dystrophin” as a key inaccuracy.*

Sarepta: *“The random measurement error of our Western blot protocol for measurement of dystrophin levels was well below the observed difference between untreated and treated Week 180 biopsy samples. A rigorous validation of the Western blot method was reviewed by the FDA prior to Week 180 biopsy analysis. Validation data demonstrated a %CV of +/- 50% and a linear range (R2 >0.9) of sensitivity extending as low as 0.25% of normal.”*

FDA response: *As quoted above, prior to analysis of the 4th biopsy, FDA explained to the applicant that major sources of random error were the results of both intra- and inter-individual differences, including differences in dystrophin that might occur within the same muscle group, or even within different regions of a single biopsy sample.¹¹ The applicant’s discussion of the variability of the Western blot method does not consider these potentially large sources of biological variability.*

¹⁰ The applicant notes that Week 180 samples were measured relative to a single normal individual’s deltoid muscle biopsy, which introduces additional uncertainty into the interpretation of fold increase vs. normal because dystrophin appears to vary about 2-fold among different normal individuals.

¹¹ Anthony et al (2014) Dystrophin quantification, biological and translational research implications. Neurology 83:1-8.

Percent Positive Fibers

Table 1 shows the percent positive fibers in eteplirsen patients. On average, the percentage of fibers with any detectable staining was about 17%, versus about 1% in the controls selected by the applicant. It is important to stress, however, that the applicant's definition of a positive fiber was *not* based on a threshold amount of dystrophin or staining brightness, but rather only on "a majority of the fiber perimeter stain at an intensity judged by eye to be *above background* of the image." [emphasis added] Consequently, "17% positive fibers" does not correspond to 17% of normal dystrophin levels, or to 17% of fibers being as bright as in BMD. The percent positive fiber result is, instead, mainly useful for localization of dystrophin, not quantification.

It is important to stress that 17% positive fibers does not represent 17-times more dystrophin compared 1% positive fibers, and is consistent with the estimate of 0.9% of normal dystrophin from Western blot. Most fibers counted as positive were faintly stained. The amount of dystrophin per fiber that would correspond to this faint immunofluorescence is unknown, but if it were 5% of normal, then 17% positive fibers with each fiber containing 5% of the normal level of dystrophin would contain $17\% \times 5\% = 0.85\%$ of normal levels of dystrophin, essentially the same value that was obtained by Western blot.

For dystrophin levels above the applicant's lower limit of reliable detection for Western blot, 0.25%, there was little correlation between Western blot and percent positive fibers, although the extent to which this represents a true inconsistency vs. random noise is not clear.

Table 1: 4th Biopsy Western Blot and %Positive Fibers, Eteplirsen Patients

Patient	Western Blot	% Positive Fibers
A	2.05	18.5
B	1.15	19.1
C	0.38	33.5
D	1.62	24
E	0.52	21.5
F	0.98	12.8
G	0	7.1
H	2.47	20.7
I	0.96	28.2
J	0	1.4
N	0.14	4.5

Total Dystrophin Immunofluorescence Intensity

There was about a 2-fold increase in overall immunofluorescence intensity in tissue sections as measured by semi-quantitative immunofluorescence (Bioquant). As discussed below (Section f), there is no simple or reliable way to compare estimates of dystrophin amount derived from overall immunofluorescence with estimates derived from Western blot.

Note: *Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statement, “There is no simple or reliable way to compare estimates of dystrophin amount derived from immunofluorescence with estimates derived from Western blot” as a key inaccuracy.*

Sarepta: *“Correlation between dystrophin quantification by Western blot and IHC methods has been demonstrated by multiple laboratories (Taylor, 2012; Anthony, 2011; Anthony, 2014; Hathout, 2015 FDA Workshop on Measuring Dystrophin).”*

FDA response: *WB is generally the more reliable method for dystrophin quantification, whereas IF is used primarily for localization of dystrophin. WB data is available, such that the strength of correlation between dystrophin quantification by the two methods is not a key issue for understanding whether or how much dystrophin may be produced by eteplirsen. Regarding the specific work cited by the applicant, the correlation between IF and WB is higher at dystrophin levels that are above those encountered in eteplirsen studies; however, the correlation is low at the low levels of dystrophin in eteplirsen treated patients.*

Importantly, the applicant digitally altered¹² dystrophin images in their background material (images in Appendix 12) such that low intensity values were increased to produce a higher intensity and higher contrast image. We are concerned that this type of image alteration makes dystrophin levels appear closer to those of BMD patients than they truly are.

¹² Per the applicant: “To generate the enhanced inverted_b base100 Image (InvertBase100), the algorithm produces a non-linear mapping of r,g,b fluorescent values that will specifically enhance low contrast objects in the image. It does this by scaling the r,g,b fluorescent values using the following formula: $I' = 1 - 100^{(-I)}$ normalized by the max value of $1 - 100^{(-1)}$ for each of the channels independently. This results in low intensity values being stretched and therefore perceived as having a higher intensity and a higher contrast”

e. Dystrophin in BMD

Quantity: The minimum level of Becker-type dystrophin that might be reasonably likely to predict clinical benefit remains unknown, but experts in DMD,¹³ including those directly involved in the development of eteplirsen,¹⁴ have stated that levels less than 3% of that of normal healthy muscle, as identified by Western blotting, are generally associated with the typical DMD phenotype, and have proposed, based on a wide range of scientific observations, that “induction of approximately 10% of normal dystrophin levels sets a minimum level to confer measurable clinical benefit.”¹⁵

Dystrophin levels in exon-51 model BMD patients have been observed to be much higher than these estimates, roughly 80% of normal on average.¹⁶ The clinical phenotype in these patients is, however, generally much milder than DMD, and this should not be taken to suggest that such high levels would be necessary for any benefit.

Since the discovery of revertant fibers and trace dystrophin in DMD, investigators have looked for, but generally not found,¹⁷ a correlation between DMD severity and trace levels of dystrophin. However, interpretation of studies is limited by questions of reliability and comparability of methods, and lack of consistent and quantitative definition of “trace” or “low level” dystrophin. For example, in one report that found a relationship between low levels of dystrophin and clinical severity of DMD, the dystrophin levels that correlated with a milder course appeared to be substantially higher than 3%,¹⁸ perhaps 15%, as measured by Western blot. Another report failed to find a correlation between the presence of reverted fibers and the clinical severity of DMD, and found a less severe clinical course only in a limited number of patients

¹³ Flanigan KM (2014) Duchenne and Becker muscular dystrophies. *Neurol Clin.* 32,1 671-688.

¹⁴ Lu QL, Cirak S, Partridge T (2014) What can we learn from clinical trials of exon skipping for DMD? *Mol Ther Nucleic acids.* 3, e152.

¹⁵ Wilton SD, Veedu RN, Fletcher S (2015) The emperor’s new dystrophin: finding sense in the noise. *Trends in Molecular Medicine.* 21, 417-426.

¹⁶ Anthony K et al (2011) Dystrophin quantification and clinical correlations in Becker muscular dystrophy: implications for clinical trials. *Brain.* 134, 3544-3556.

¹⁷ Flanigan KM (2014) Duchenne and Becker muscular dystrophies. *Neurol Clin* 32:671-688.

¹⁸ Nicholson, LVB (1993) Functional significance of dystrophin positive fibers in Duchenne muscular dystrophy. *Archives of Diseases in Childhood.* 68:632-636.

showing a faint dystrophin labeling in most fibers.¹⁹ Patients who are amenable to exon 44 skipping have been reported to express higher levels of dystrophin than in DMD patients with other exon-skippable mutations, and to have a somewhat milder course, but it is not clear how much dystrophin is expressed in these patients (most reports have focused on immunofluorescence rather than Western blot²⁰) or in what percentage of fibers (staining in nearly 100% of fibers occurs in at least some exon 44 skippable patients²¹). Possible differences in functionality of the truncated dystrophin species produced in patients with different mutations also confounds interpretation of possible effects on clinical course of differences in dystrophin levels.

Timing: Experts have cautioned that dystrophin is present in BMD from birth, and that “we should not conclude that dystrophin restitution in DMD patients with established dystrophic pathology will confer comparable benefits to the dystrophins in BMD patients”²² for reasons including the pro-inflammatory environment that develops in DMD.²³

Functionality: The exact dystrophin mutation affects the clinical phenotype in BMD,²⁴ and likely also in DMD, confounding interpretation of any possible clinical impact of small differences in dystrophin levels among DMD patients, with experts stressing that “it will be essential to account for different mutations when looking at other possible contributing factors to disease severity.”²⁵

¹⁹ Fanin et al (1995) Dystrophin-positive fibers in Duchenne dystrophy: origin and correlation to clinical course. *Muscle and Nerve*. 18:1115-1120.

²⁰ Anthony K, et al (2014) Biochemical characterization of patients with in-frame or out-of-frame DMD deletions pertinent to exon 44 or 45 skipping. *JAMA Neurol*. 71:32—40.

²¹ Beekman et al (2014) A sensitive, reproducible and objective immunofluorescence analysis method of dystrophin in individual fibers in samples from patients with Duchenne muscular dystrophy. *PLOS ONE*. 9:e107494.

²² Wilton SD, Fletcher S, Flanigan KM(2014) Dystrophin as a therapeutic biomarker: Are we ignoring data from the past? *Neuromuscular Disorder*. 24, 463-466.

²³ Rosenberg et al (2015) Immune-mediated pathology in Duchenne muscular dystrophy. *Sci Transl Med* 7,299rv4.

²⁴ Nicolas et al (2015) Becker muscular dystrophy severity is linked to the structure of dystrophin. *Human Molecular Genetics*. 24:1267-1279.

²⁵ Van den Bergen JC et al (2014) Dystrophin levels and clinical severity in Becker muscular dystrophy patients. *Neurol Neurosurg Psychiatry*. 85, 747-753.

Localization: In BMD, dystrophin is typically present in all or most fibers^{26,27} and, in addition to the total amount, this is thought to be important for function of the dystrophin. In contrast, in DMD many patients have no detectable dystrophin staining, while others have bright staining in a small percentage (1- to 5%) of “revertant” fibers in which exon skipping is thought to occur spontaneously. Some DMD patients can also show faint dystrophin staining in up to about 25% of fibers,²⁸ with the percentage of positive fibers appearing to depend in part on technical factors that affect assay sensitivity.

Low level dystrophin immunofluorescence in almost 100% fibers has also been reported in DMD, including in exon-51 skippable patients.²⁹

Unusual BMD Patients: Rarely, patients with BMD are encountered who have dystrophin levels that are less than 1% of normal, which is as low as typical DMD patients. Importantly, however, rather than suggesting that very low levels of drug-induced dystrophin are likely to be beneficial, such patients highlight the complexity of the relationship between dystrophin levels and phenotype. The fact that such patients can have mild disease appears to be unrelated to, not necessarily the result of, low levels of dystrophin. In this context, the applicant selected three BMD patients as comparators for the Week 180 dystrophin studies, one of whom had low dystrophin level of about 2% of normal. However, the BMD patients selected by the applicant do not appear representative, and this patient may correspond to one of the rare BMD patients with very low dystrophin levels.

***Note:** Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statements, “In this context, the applicant selected three BMD patients as comparators for the Week 180 dystrophin studies, one of whom had low dystrophin level of about 2% of normal. However, the BMD patients selected by the applicant do not appear representative,*

²⁶ Arahata et al (1989) Dystrophin diagnosis: comparison of dystrophin abnormalities by immunofluorescence and immunoblot analysis. Proc Natl Acad Sci. 86,7154-7158.

²⁷ Morandi et al (1995) Dystrophin characterization in BMD patients: correlation of abnormal protein with clinical phenotype. Journal of Neurological Sciences 132,146-155.

²⁸ Arechavala-Gomez et al (2010) Revertant fibres and dystrophin traces in Duchenne muscular dystrophy: implications for clinical trials. Neuromuscul Disord. 20,295-301.

²⁹ Beekman et al (2014) A sensitive, reproducible and objective immunofluorescence analysis method of dystrophin in individual fibers in samples from patients with Duchenne muscular dystrophy. PLOS ONE. 9:e107494.

and this patient may correspond to one of the rare BMD patients with very low dystrophin levels.” as a key inaccuracy.

Sarepta: *“BMD patient samples were not chosen to be representative; rather, they were selected in response to an FDA request to assess the relationship between dystrophin as measured by Western blot and immunofluorescence fiber intensity. Therefore, BMD samples were obtained that represented low, middle, and higher ranges of dystrophin expression. A comparable Western blot analysis - IHC correlation was presented by Hathout, et al. (MDA 2015 Scientific Conference poster, FDA - NIH workshop on measuring dystrophin, 2015), where BMD biopsies were chosen to represent low- and mid-level dystrophin expression. Consistently, their BMD low patient biopsy was 2% of normal.”*

FDA response: *It isn’t clear that there is any disagreement. The BMD patient selected by the applicant, who has dystrophin levels of about 2% of normal, is not representative of levels typically associated with BMD, and may correspond to one of the rare patients whose clinical course is milder than expected despite low levels of dystrophin typically associated with the DMD phenotype.*

As further illustration, there are rare cases of siblings where both show a negative pattern of dystrophin immunostaining and scattered revertant fibers yet have highly discordant phenotypes. For example, Zatz et al³⁰ reported a case of nonsense mutation DMD in which the younger brother was wheelchair-bound at age 9 years, whereas his half-brother was reported to have some difficulties running and climbing stairs at age 15 years but normal walking ability.

f. Reviewer Discussion, Dystrophin Quantification Methods

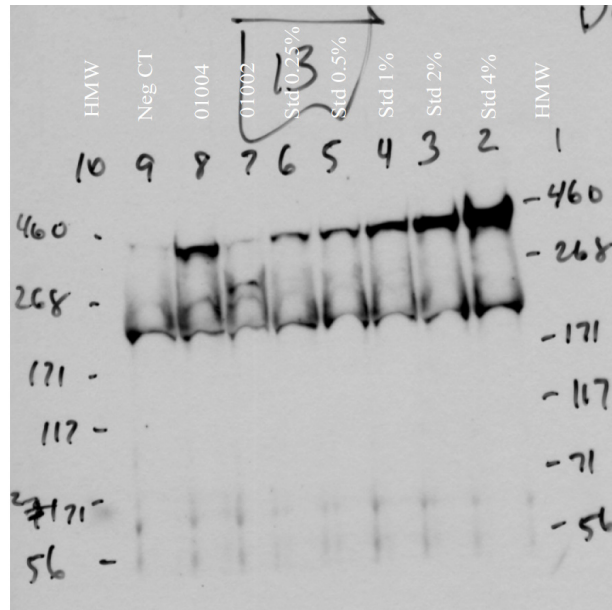
Considerable confusion can be created by the fact that a number of different methods have been used to quantify dystrophin expression, some more quantitative than others, and some producing higher absolute numbers than others. As discussed above, immunofluorescence is mainly informative of dystrophin localization, but is not a reliable measure of dystrophin amount (beyond perhaps the binary distinction between “undetectable” and “detectable”). For example, in many patients with typical DMD, only trace levels of dystrophin are present, yet these levels result in 25% or more of fibers being faintly dystrophin-positive.

Western blot, in contrast, cannot provide information about dystrophin localization within the tissue, but does allow reasonable quantification through the use of internal controls with

³⁰ Zatz M et al (2014) Milder course in Duchenne patients with nonsense mutations and no muscle dystrophy. *Neuromuscular Disorders*. 24:986-989.

defined amounts of dystrophin (currently defined in terms of percent of dystrophin of a normal individual, not purified protein, which does introduce a small amount of uncertainty, but perhaps 2-fold or less). A dilution series control is shown in Figure 4, near the “460” molecular weight marker, from right to left.

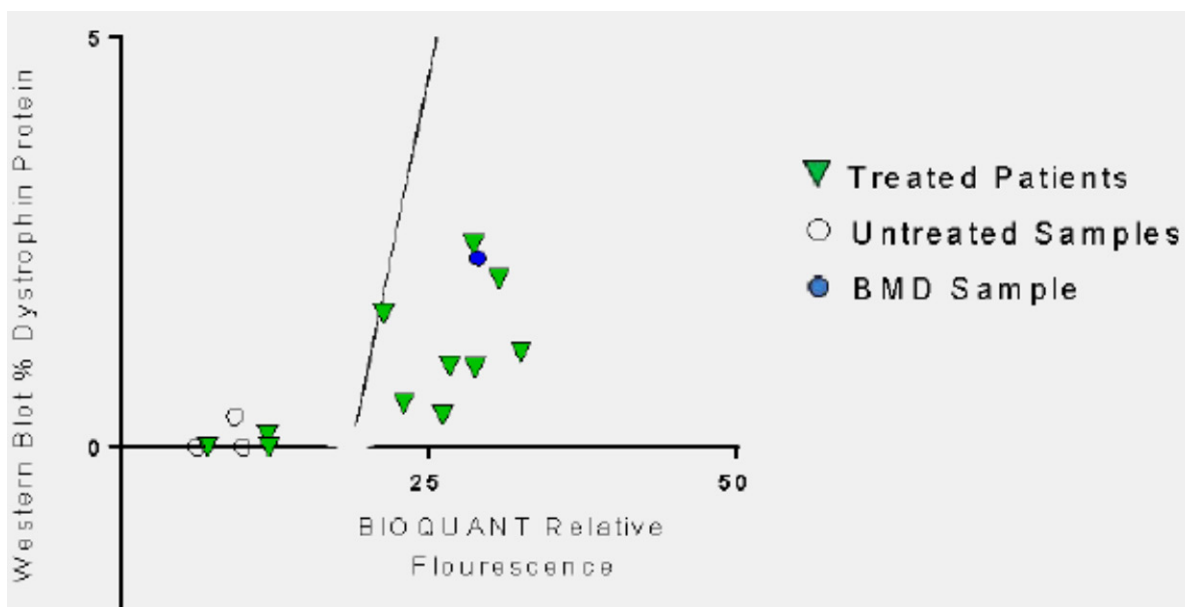
Figure 4: Western blot, 4th Biopsy, Study 202



In contrast, immunofluorescence methods lack similar internal controls, and as a consequence it is essentially impossible to correlate a certain amount of fluorescence to a certain amount of protein measured by Western blot, or relative to a normal control. There is no simple or reliable way to compare estimates of dystrophin amount derived from immunofluorescence with estimates derived from Western blot.

Figure 5 shows that at low levels of dystrophin (<5% by Western blot), immunofluorescence appears to overestimate the amount of dystrophin; for example, immunofluorescence shows about 25% intensity for samples with roughly 1- or 2% of normal dystrophin by Western blot, and shows about 10% of normal intensity for samples with <1% of normal dystrophin levels.

Figure 5: Western blot vs. Bioquant



Finally, a representation of the change in dystrophin levels in terms of percent change from baseline is problematic in this situation, because the trace baseline dystrophin levels in many patients are too low to be measured accurately, resulting in ratios that are imprecise, and that are greatly affected by small amounts of random variability in denominators that are close to zero.

Expressing dystrophin levels as percent- or fold-change compared to controls exaggerates the difference:

- *Dystrophin levels that were, in fact, detected but that were less than 0.25% were imputed as zero.*
- *The lower limit of reliable detection of the assay is 0.25%. It would be more accurate to consider undetectable dystrophin levels as <0.25%, not as zero.*

g. FDA Review Team Preliminary Conclusions on Dystrophin Findings

Adequate scientific methods appear to be available to measure dystrophin expression in DMD. As discussed in the recent FDA draft Guidance on DMD,³¹ there is justifiable interest in dystrophin as a potential surrogate endpoint for accelerated approval in DMD. However, the

³¹ Duchenne muscular dystrophy and related dystrophinopathies: developing drugs for treatment. <http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>

Guidance also states that the potential for a biomarker to predict clinical benefit in DMD is inseparable from such factors as the magnitude of change of the biomarker. Regarding methodology, the Guidance stresses the importance of the performance characteristics of the biomarker assays, including quality-control measures.

Based on the data submitted by the applicant, considerable doubt remains about how much, or perhaps even whether, dystrophin levels were increased by eteplirsen. The degree of uncertainty about the dystrophin data hinders discussion of its use as surrogate endpoint for eteplirsen. However, to the degree that the dystrophin data may be interpretable, the amount and distribution of dystrophin in treated patients appears to be within the range typically associated with DMD, not BMD. Data suggesting that higher levels of dystrophin were produced by eteplirsen appear unreliable.

Clinical Efficacy Evidence

The only study that evaluated clinical efficacy is Study 201/202. Dr. Xiang Ling, from the Office of Biometrics, provided a statistical review of that study. As described below, and in Dr. Ling's review, Study 201/202 was not designed in a way that allows reliable use of statistical hypothesis testing (i.e., "p-values"), and is only capable of providing interpretable evidence of efficacy if the beneficial effect of eteplirsen is so large that it is essentially self-evident, without the use of statistics.

a. Design and analysis of Study 201/202

Clinical efficacy was examined in one single-center, 24-week, 3-arm controlled trial (Study 201) in 12 patients assigned 1:1:1 to 30 mg/kg eteplirsen, 50 mg/kg eteplirsen, or placebo. Study 201 was continued as an open-label extension, called Study 202, which has been ongoing for more than 3 years. Multiple functional endpoints were assessed both in the placebo-controlled and open-label extension periods, including 6 minute walk test (6MWT), North Star Ambulatory Assessment (NSAA), and a number of measures of pulmonary function. Analysis of clinical endpoints was not controlled for multiplicity, but in Study 201 the clinical endpoints were essentially uniformly negative, without trends supportive of efficacy.

Note: *Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statement in the statistical review that "the robustness of the study result is a concern since a single patient could change the results substantially"*

Sarepta: *“This statement is inaccurate. A comprehensive sensitivity analysis was performed in order to address any potential issue regarding robustness of the data. Specifically:*

- *Two patients were removed: the best performing eteplirsen and the worst performing external control patient.*
- *Results demonstrated a robust 6MWT treatment advantage of >100 meters with nominal significance.”*

FDA Response: *This statement from the statistical review appears to be in reference to the placebo-controlled portion of Study 201/202, which was small in size (N = 4 per arm), such that changes in the outcome measure for a single patient could change the overall results substantially. The statistical review also notes that a key limitation of the externally controlled open-label portion of Study 201/202 was dissimilarity of the groups being compared, along with differences in how the data were collected, as also detailed in this memo and other background information from the FDA. The applicant’s statistical approach to analysis of the externally-controlled portion of Study 201/202 does not address the key source of uncertainty in any externally-controlled trial: the presence of non-drug related differences between groups, some of which are known, and some of which are unknown. One of the applicant’s proposed sensitivity analyses, which removed the single best-performing eteplirsen patient and the single worst performing external control patient, does not address this fundamental issue.*

Shortly after Study 202 passed 1 year duration, the applicant proposed a post-hoc analysis with a number of changes from the original analysis: a) data for 2 out of 8 patients treated with eteplirsen (patients who quickly lost ambulation) were dropped, b) the prespecified comparison of each dose arm to placebo was changed to comparison of the 6 remaining treated patients to the 4 placebo-treated patients, and c) the endpoint was taken to be Week 36, instead of Week 24. FDA explained in detail to the applicant in March of 2013 why the proposed analysis was unreasonable even for hypothesis generation, and why Study 201 did not provide evidence of efficacy.

As the duration of exposure in Study 202 increased, the applicant proposed comparing the clinical course of treated patients to historical controls. FDA expressed strong reservations regarding the potential interpretability of the applicant’s proposed comparison to historical controls and the use of 6MWT as the primary endpoint in such a historical comparison. Because of these concerns, FDA noted that a dramatic effect size would be necessary for any such analysis to be potentially interpretable. Well-designed historically-controlled trials can, in

certain circumstances, be considered adequate and well-controlled designs that can support FDA approval. However, Study 201/202 is not a well-designed historically-controlled trial. It is well established, as detailed in guidelines developed by U.S. and international regulatory bodies,³² that “inability to control bias is the major and well-recognized limitation of externally-controlled trials, and it is always difficult, and in many cases impossible, to establish comparability of the treatment and control groups.” Furthermore “a consequence of the recognized inability to control bias is that the potential persuasiveness of findings from externally controlled trials depends on obtaining much more extreme levels of statistical significance and much larger estimated differences between treatments than would be considered necessary in concurrently controlled trials.”

Note: *Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statement, “As the duration of exposure in Study 202 increased, the applicant proposed comparing the clinical course of treated patients to historical controls” as a key inaccuracy.”*

Sarepta: *“The proposal to compare with historical control patients originated from the FDA. Specifically, a requirement to compare the clinical course of treated patients in Study 202 to matched patient-level historical control data was made by the FDA at the March 2014 guidance meeting, and reiterated at the September 2014 pre-NDA meeting. Sarepta had proposed an open-label confirmatory study comparing treated patients to concurrent (not historical) untreated patients with exon deletions not amenable to skipping exon 51 (i.e., the PROMOTI study).”*

FDA response: *FDA consistently and strongly encouraged the sponsor to conduct adequately powered randomized placebo-controlled trials, and expressed doubt about the interpretability of externally controlled trials. As early as October 2012, Sarepta and its academic associates announced that in the randomized controlled portion of Study 201/202 eteplirsen had demonstrated unparalleled effects on enabling dystrophin production and slowing the progression of the disease,³³ with levels of dystrophin potentially as high as 50% of normal. In the context of an ongoing series of reports from the applicant and its academic associates describing continued striking and*

³² Choice of control group and related issues in clinical trials, E10. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2000.

³³ <http://investorrelations.sarepta.com/phoenix.zhtml?c=64231&p=irol-newsArticle&ID=1741044>; accessed March 17, 2016.

unprecedented stabilization of disease progression, many in the DMD community expressed strong reservations regarding the ethics of conducting another placebo-controlled trial, and informed FDA that performing such a study would be extremely difficult or impossible. In this context, and based on assertions that eteplirsen had been shown unequivocally to produce high levels of dystrophin, FDA expressed willingness to consider an externally controlled trial, although stating clearly that interpretation of the data could be difficult, and that the acceptability of the study would be a matter for NDA review.

FDA informed the applicant that if it were to pursue a comparison of patients in Study 201/202 to external controls, evaluating such a comparison would be difficult without submission of patient-level external data, including data from a number of different sources to understand variability across different datasets, which can be substantial in DMD. For example, Biggar et al³⁴ reported that about 75% of a population of DMD boys treated with deflazacort was ambulant at age 15 years (N = 40), whereas Bello et al³⁵ reported that in data collected by the Cooperative International Neuromuscular Research Group (CINRG) about 25% boys³⁶ similarly treated with deflazacort were ambulatory at age 16 years (N = 80).

After release of the previous version of this memo, CINRG provided additional unpublished analyses to FDA suggesting that exon-51 skippable patients follow a clinical course for age of loss of ambulation generally similar to that described for the broader DMD population described in Bello et al, with about 25% of boys maintaining ambulation to 16 years of age and about 15% of patients maintaining ambulation to 18 years of age. At the time this revised memo was written, CINRG was in the process of providing patient-level CINRG data to FDA that should enable more detailed comparison with eteplirsen-treated patients for both age at loss of ambulation and functional endpoints such as 6MWT and 10 m walk/run, based on a prespecified plan.

³⁴ Biggar WD et al., (2006). Long-term benefits of deflazacort treatment for boys with Duchene muscular dystrophy in their second decade. *Neuromuscular Disorders* 16:249-255.

³⁵ Bello L et al (2015) Prednisone/prednisolone and deflazacort regimens in the CINRG Duchenne natural history study. *Neurology*. 85, 1048-1055.

³⁶ CINRG has subsequently provided FDA with unpublished analyses suggesting similar natural history in exon-51 skippable patients, as discussed elsewhere in this review.

Note: Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statements, “Finally, as the natural history studies proceeded, some patients left to enter interventional clinical trials, further decreasing the similarity of the natural history cohort to the eteplirsen patients” as a key inaccuracy.

Sarepta: “Two types of missing data sensitivity analyses were performed, the results confirmed that the magnitude of difference remained over 100 meters and nominal statistical significance was maintained:

- MMRM using all the available data
- Last Observation Carried Forward imputation (conservative analysis assuming that the 2 control patients did not decline)”

FDA response: It should be stressed that for a variety of reasons the clinical course of patients in recent observational studies in DMD, including CINRG, might be expected to be worse than the clinical course of patients selected for studies of experimental drugs. Differences in patient selection, supportive care, motivation, and how loss of ambulation is defined and measured, among other factors, are likely to be important.

Various analytical methods to impute missing data, such as mixed effect model repeat measurement (MMRM) and last observation carried forward (LOCF), do not address the key limitation of a comparison between an open-label treatment group in an interventional clinical trial and an independent group of patients who are in an observational study: non-drug-related differences between the groups being compared.

Recent observational studies in DMD have been enrolling patients simultaneously with interventional trials of new drugs. Thus, patients in an observational cohort who were motivated to enroll in a drug study and could qualify for enrollment might have preferentially left the observational study. In other words, patients who remained in the observational study may have been less motivated or less able to participate in studies of experimental drugs. Moreover, patients in an observational study are likely to differ in other important ways. Specific evidence of this effect appears to be present in the historical data submitted by the applicant. A patient selected as a historical control for Study 201/202 lost ambulation after a single 6MWT measure, and stayed in the observational study for several years, long enough to be matched to eteplirsen patients. In contrast, two other exon-51 patients with similar baseline age and 6MW distance discontinued the observational study to participate in drug studies. These patients,

doing reasonably well, were therefore not under observation for long enough to serve as historical controls for the eteplirsen study.

Many aspects of supportive care are important for prolonging function in DMD, yet difficult to quantify, and this appears to be particularly true for physical activity. Regular physical activity is necessary to maintain function in DMD and to avoid disuse atrophy.³⁷ Gentle exercise appears to provide additional benefit, including delay of functional deterioration.³⁸ Use of a wheelchair may justifiably be encouraged by caregivers for reasons of safety and independence, or even be required in settings such as school. In addition, although difficult to quantify, accounts by caregivers suggest that pessimism and resignation about prognosis in DMD may contribute to decreased time spent walking and less independent activities and self-care, whereas feelings of hope and optimism from enrolling in a drug study may lead to the opposite behavior. Particularly in muscular dystrophy, it therefore seems possible that hope and positive expectations might increase physical activity and decrease the risk of disuse atrophy, thus slowing functional decline. Slower decline or even improvement in function have been observed in placebo arms of controlled trials in other types of muscular dystrophy,³⁹ and potentially may be the result of some of the above mechanisms.

FDA encouraged the sponsor at the March 2013 meeting to conduct an adequately powered placebo-controlled trial of eteplirsen, stating “if it is true that eteplirsen leads to remarkable clinical benefit in even some patients, there is no doubt that a feasible placebo controlled study can be designed to demonstrate that benefit.” FDA also stated that “there is considerable variation among individual patients with regard to clinical measures and important milestones” and that data from an open-label study “may only be interpretable if a relevant objective endpoint obviously insulated from bias demonstrated compelling data that are clearly outside the known variability range for DMD.” FDA further stated that, at that time, comparison of data from Study 202 did not provide interpretable evidence of benefit “given the limitations of the open-label design for protecting against bias on effort-dependent endpoints like 6MWT.” At a July 2013 meeting with the applicant, at which the possibility of NDA filing based on dystrophin

³⁷ Bushby K et al (2009) Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. The Lancet. DOI:10.1016/S1474-4422(09)70271-6

³⁸ Jansen M et al (2013) Assisted bicycle training delays functional deterioration in boys with Duchenne muscular dystrophy: the randomized controlled trial “no use is disuse”. Neurorehabil Neural Repair. DOI: 10.1177/1545968313496326

³⁹ Statland JM et al (2013) Reevaluating measures of disease progression in facioscapulohumeral muscular dystrophy. Neuromuscul Disord. 23,306-12.

production was discussed, FDA similarly expressed reservations about natural history controls “due to the usual difficulty in showing comparability between the study populations in natural history studies,” and reiterated that 6MWT was susceptible to bias in the proposed natural history comparison.

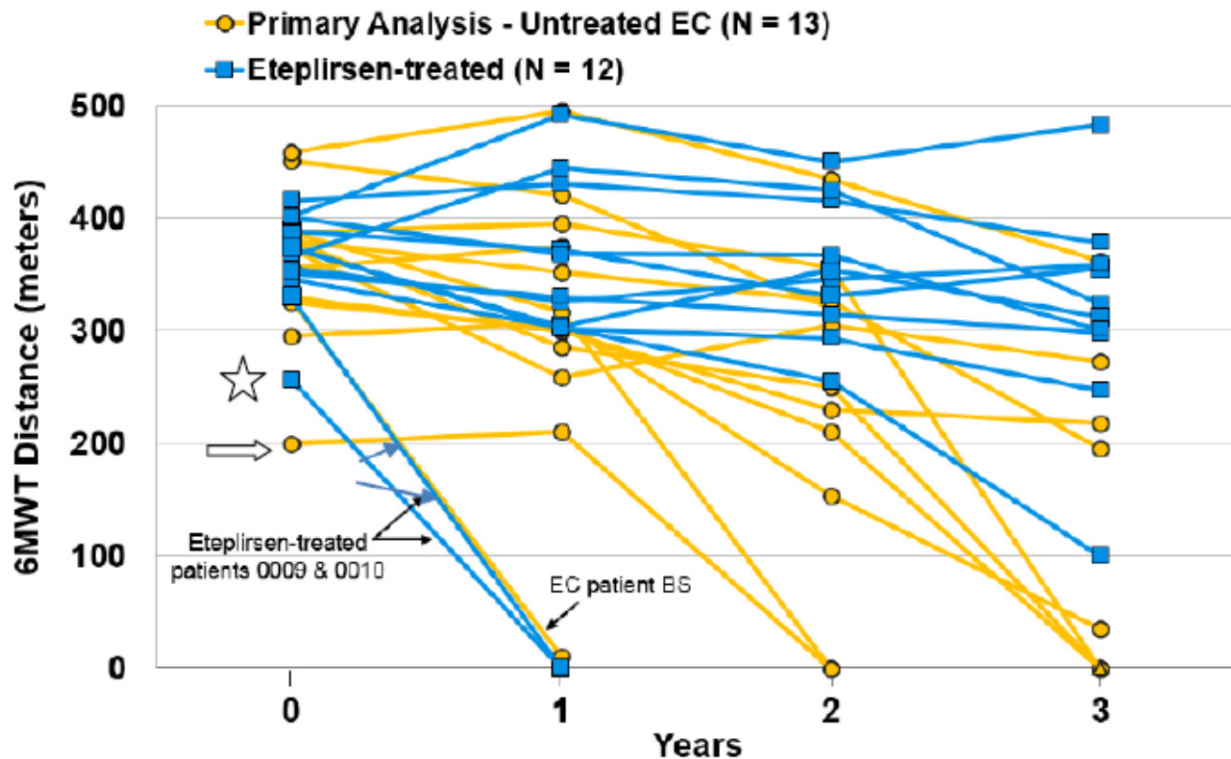
Discussions about comparison of Study 202 patients to natural history continued with the April 15, 2014, communication from FDA to the applicant which stated that, with additional data to support the efficacy and safety of eteplirsen, an NDA should be filable. FDA noted that patients in Study 202 appeared to be receiving optimal care, including intensive physical therapy and intensive steroid regimens, and again stated that “performance on the 6-minute walk test is strongly influenced by motivation and coaching, and open-label trials are susceptible to bias on the part of investigators, patients, and parents.” In a September 2014 communication, FDA explained its concern that, as noted by DMD experts, “preservation of ambulation and other skills is affected by the value that families and caregivers put on maintaining those skills, with such factors as risk of falls and injury from continued ambulation weighed against the safety and speed of allowing patients to use a wheelchair.” FDA further advised the applicant that it was not clear that such biases could be adequately controlled, and that the applicant should present data from measures of muscle strength in the NDA to assist in determining if measures of ambulation had been affected by these types of bias. As discussed below, results from rise time measures and the NSAA appear to be reasonable measures of muscle strength in this context, and thus important for interpreting the 6MWT results.

To interpret the applicant’s comparison of 6MWT results for eteplirsen patients to historical controls, it is also important to understand the progression of 6MWT as DMD patients near the time of loss of ambulation. At younger ages, during the period of relative stability or slow decline of 6MWT, a difference between two patients in 6MWT of 100 m is likely to predict a difference of several years in time to loss of ambulation, particularly if one patient is below about 300 meters and the other above. Differences between patients of 150- or 200 m on 6MWT have even larger prognostic implications, with patients who can walk in the range of 400- to 500 m on 6MWT unlikely to lose ambulation for many years. In contrast, however, large differences in 6MWT between patients near the time of loss of ambulation occur even when patients have generally similar prognoses.

Figure 6, taken from the applicant’s NDA, shows patient-level data for eteplirsen and historical controls. Consider two patients in their final year or two of ambulation: the historical control patient with a baseline of about 200 m (arrow), and the eteplirsen patient with a baseline of about 260 (star). At Month 12, the eteplirsen patient has lost ambulation, whereas the 6MWT for the historical control patient remains at about 200 m, such that the difference in 6MWT has

increased from 60 m at baseline to about 200 m. By Month 24, the historical control patient has also lost ambulation, such that the difference between patients has become zero. Thus, in contrast to younger patients, the 200 m difference near the time of loss of ambulation corresponded to about 1 year difference in age at loss of ambulation. The general pattern and size of this effect is typical, with many DMD patients decreasing from about 300 m on 6MWT to loss of ambulation over 1- to 2 years, leading to brief but very large differences in 6MWT between patients whose disease course is otherwise generally similar. This does not imply that a difference of 150- or 200 m on 6MWT would not be clinically meaningful, but does suggest that even modest differences between study arms in poorly controlled studies such as Study 202 can exaggerate differences in certain functional measures near the time that patients lose ambulation.

Figure 6: 6MWT in Patients Using Steroid, Age ≥ 7 Years, Amenable to Exon 51 Skipping by Treatment Status – Individual Patient Data



b. Rate of progression of 6MWT in eteplirsen-treated patients is consistent with expected natural history

Data reliability is a major concern in the comparison of eteplirsen-treated patients from Study 201/202 to external controls. It has been suggested to FDA by a number of outside individuals and groups that ambulation is a reliable efficacy endpoint in historically-controlled trials in DMD because it is a “hard” endpoint, i.e., an objective, invariant state indicating inability to walk independently. However, near the time of loss of ambulation factors such as effort and motivation on the part of both patient and examiner can have very large effects on ambulatory endpoints, such that loss of ambulation cannot be considered a “hard” endpoint in this setting. A 6-minute walk distance of 0 meters, or isolated or even consecutive zero values resulting, for example, from an injury from which the patient recovers, does not necessarily represent irreversible inability to walk.

Subsequent to the release of the previous version of this memo, FDA has determined that for at least two or three⁴⁰ of the 13 exon-51 skippable natural history patients selected by the applicant as controls, a value of zero was recorded for 6-minute walk distance apparently prior to loss of ambulation as documented by ability to perform the 10 meter walk/run test. Similar discordance between 6MW distance and 10 m walk/run was identified for at least 6 patients in the group of external control patients. Importantly, for both the exon-51 skippable patients and larger group of external controls, 10 m walk/run data were not available for many patients, limiting ability to assess discordance of results.

- At age 12, one exon-51 skippable control patient from Belgium was recorded as having a 6MW distance of 327 m, and a 10 m walk time of 7 s. At the next exam about 6 months later, 6MW distance was recorded as zero, but the patient was able to complete the 10 m walk in 11 s. This pattern continued with the next two exams over the following year, with 10 m walk values of 11s and 13 s, yet a 6MW distance of zero.*

The applicant has recently provided FDA with source documents from the clinical sites for this patient and the other historical controls. These documents appear to indicate that at a follow-up visit 6 months later, 6MWT was not attempted because the patient was judged to be unable to walk. At the next visit 6 months later (1 year

⁴⁰ An additional exon-51 skippable patient had a 10 m walk time of 35s, and 6MWT of zero. Under some conventions, 6MWT would not be measured if the 10 m walk time is >25s, but it is not clear that consistent conventions were adopted across the natural history studies and Study 201/202.

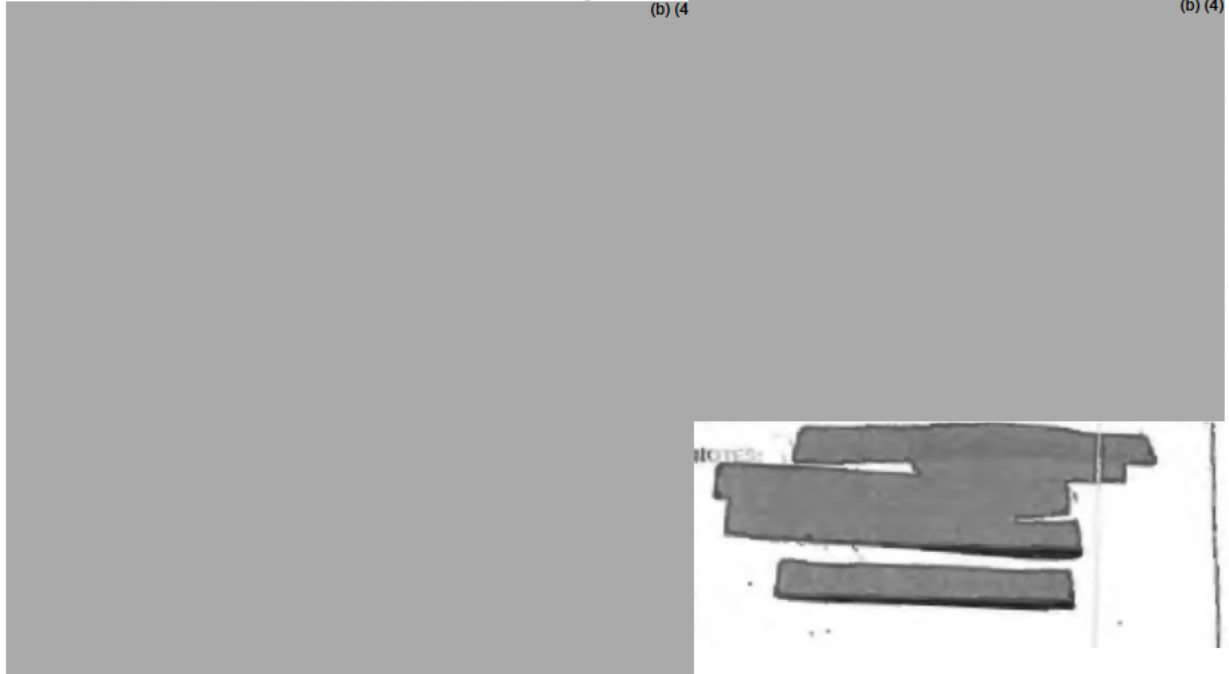
after the 327 m was recorded), a 6MWT was attempted, with the patient walking 125 m in about 3½ minutes. The examiner at the time noted that the patient “no longer wanted to continue (could still continue, had back pain).” The examiner’s comment appears to underscore the importance of motivation in 6MWT.

- *At age 10, one exon-51 skippable control patient from Italy was recorded as having a 6MWT of 356 m, and a 10 m walk time of 10 s. One year later, at age 11, 10 m walk/run time was 12 s, but 6MWT was apparently not attempted and was recorded as zero (source documents state “not executable”).*

Similar concern about reliability exists for 3 additional⁴¹ exon-51 skippable natural history patients for whom 6MW distance was reported as zero but apparently not measured. Initial review of source documents recently received by FDA suggests the applicant asked the investigators in December 2015 if patients who had been last recorded in the clinic several years previously had maintained ambulation 4 years post-baseline.

There are, in addition, low 6MW distance values recorded for natural history controls that appear atypical for reasons that are not well documented. The image below shows a source document from a historical control patient who walked for only about 1½ minutes during the 6 minute test, and was recorded as having a final distance of 35 m (note: 50 m appears to have been the total distance, but due to an apparent error the value for “1 minute distance” of 35 m was transcribed). The notes section appears to have been blackened out. For other patients, this section of the document contained important information about patient performance during the test, such as “good cooperation” or the number of times that the patient paused walking during the test.

⁴¹ *Per applicant addendum for February 23, 2016 AC meeting: “patients MI15 and SL were subsequently reported to have loss of ambulation with “0” meters on the 6MWT at ~4.5 years. In addition, external control patient PV12 was known to have lost ambulation with a 6MWT of “0” at 4.8 years”*

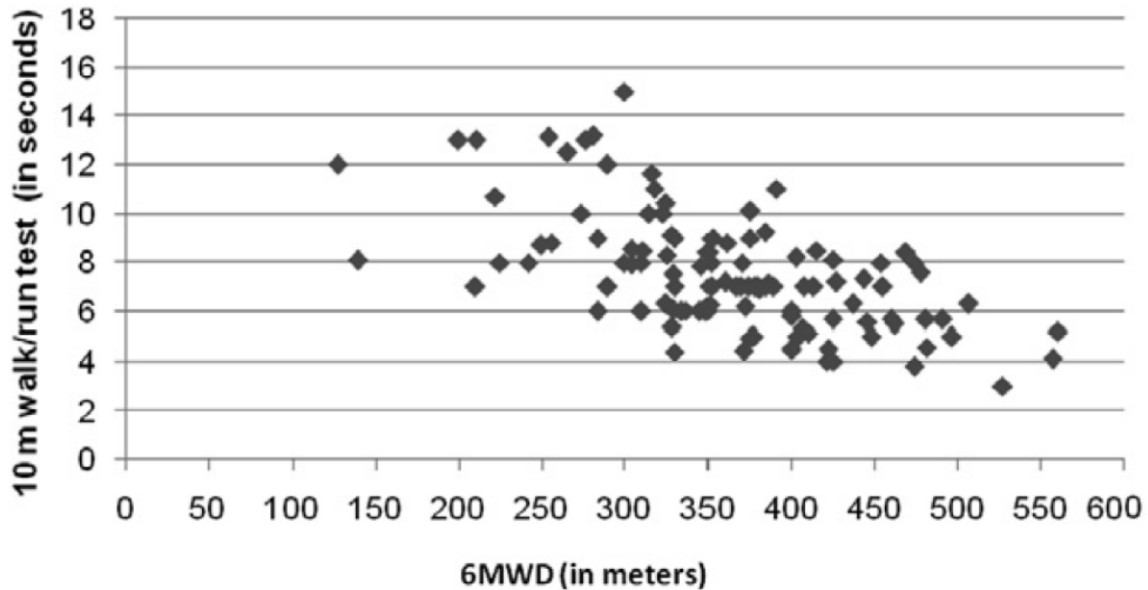


It should also be noted that eteplirsen-treated patients had two opportunities on consecutive visit days to perform functional tests, whereas natural history patients had only one. This systematic difference speaks to the dissimilarity in how the patients were managed and the level of attention given to the 6MWT in the eteplirsen study.

Datasets from the natural history studies and the eteplirsen study were examined in more detail to characterize the typical relationship between 6MWT and 10 m walk/run values that might have been expected for control patients. The investigators for the Italian natural history cohort previously reported⁴² an average 6MW distance of approximately 150 to 375 m for DMD patients with 10 m walk/run values between 11 and 13 s (Figure 7).

⁴² Mazzone et al (2010) North Star ambulatory assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscular Disorders*. 20,712-716.

Figure 7: 10 m walk/run vs 6MWD, by individual patient, Italian natural history cohort



There appeared to be a generally similar relationship between 6MWT and 10 m walk/run in eteplirsen-treated patients, for example, with values of 11 s to 12 s on 10 m walk/run corresponding to roughly 200 to 300 m on 6MWT, and 13 s to 15 s corresponding to roughly 150 to 200 m. One patient who walked 50 m on 6MWT had a 10 m walk/run time of 20 s.

Patients from the placebo arm of randomized double-blind trials are likely to be better matched to patients in eteplirsen trials for factors that are difficult to measure, such as motivation and compliance with supportive therapy, compared to patients from registries. Placebo-controlled trials have recently been conducted with patients with DMD amenable to exon-51 skipping. Data from patients from the placebo group from some of these studies are publically available, and were used for a comparison with eteplirsen-treated patients.⁴³ The figures below show the clinical course on 6MWT of eteplirsen-treated patients from Study 201/202 (colored lines) compared to patients treated with placebo in other controlled studies in exon-51 skippable patients with DMD (grey lines). Patients are divided by baseline rise from floor time (an important prognostic variable), and by steroid treatment (deflazacort, Figure 8), or prednisone

⁴³

<http://www.fda.gov/downloads/AdvisoryCommittees/CommitteesMeetingMaterials/Drugs/PeripheralandCentralNervousSystemDrugsAdvisoryCommittee/UCM475956.pdf>

(Figure 9), because some evidence suggests deflazacort may be more effective than prednisone at preserving ambulation in DMD.

A few observations about these data follow:

- Clinicians expert in the care of DMD patients often perceive that, even in patients treated with corticosteroids, decline of 6MWT after about age 7 is steady, and that periods of stability or improvement, particularly after periods of decline, do not occur. However, the placebo data show that while decline ultimately occurs, many exon-51 patients experience periods of stability or even substantial improvement. This occurs in patients older than 10 years of age, and in patients who, at least as measured by 6-minute walk distance, have experienced substantial earlier declines. This complicates the interpretation of treatment trials in DMD that may not be well-controlled.
- The figures below divide patients by baseline rise time and steroid treatment,⁴⁴ but each can be interpreted as a continuum of disease progression, from top to bottom, because the loss of ambulatory ability in DMD almost always proceeds in sequence, with rise time steadily worsening (increasing), followed by loss of ability to rise from the floor but retained ability to walk, then loss of ability to walk, which often occurs with a sharp decline when 6MWT decreases below about 300 m. Thus, even though each placebo patient was followed for only 1 year, whereas eteplirsen patients were followed for more than 3 years, there can be reasonable confidence that most placebo patients would follow a stepwise progression through higher rise times prior to loss of ambulation, such that their clinical course can be extrapolated beyond the 1 year period of observation.
- The course of 6MWT for eteplirsen patients was generally similar to the course of placebo patients across all rise time categories, and for both types of corticosteroid, with some of the placebo patients having higher (better) 6MWT than matched eteplirsen patients, and some worse. This appears to be expected given the known wide variability of progression in exon-51 DMD, and the small numbers of patients available for comparison.
- Finally, decline in 6MWT is also a reliable predictor of loss of ambulation. At the most recent study visit, 6MWT was less than 250 m for the 7 out of 10 eteplirsen patients who had maintained ambulation past the first months of the study, which also predicts a high probability of loss of ambulation in a timeframe of 1 to 2 years.

⁴⁴ Patient 7 was switched from prednisone to deflazacort in 2013, and is shown in the prednisone figure

In the figures below, many of the eteplirsen patients appear to have few or no matches to the placebo patients in the most recent year of treatment, but this is a result of the division of the figures into categories based on baseline rise time. Most eteplirsen patients are currently in the >15 s rise time category (10 of the 12 eteplirsen patients, including at least 5 who lost ability to rise), and can be compared to the >15 s rise time group of control patients. In general, the course of eteplirsen-treated patients in Study 201/202 is similar to the course in these control patients, as shown in Figure 10, which combines all eteplirsen and control patients.

Figure 8: 6MWT, Deflazacort-treated patients

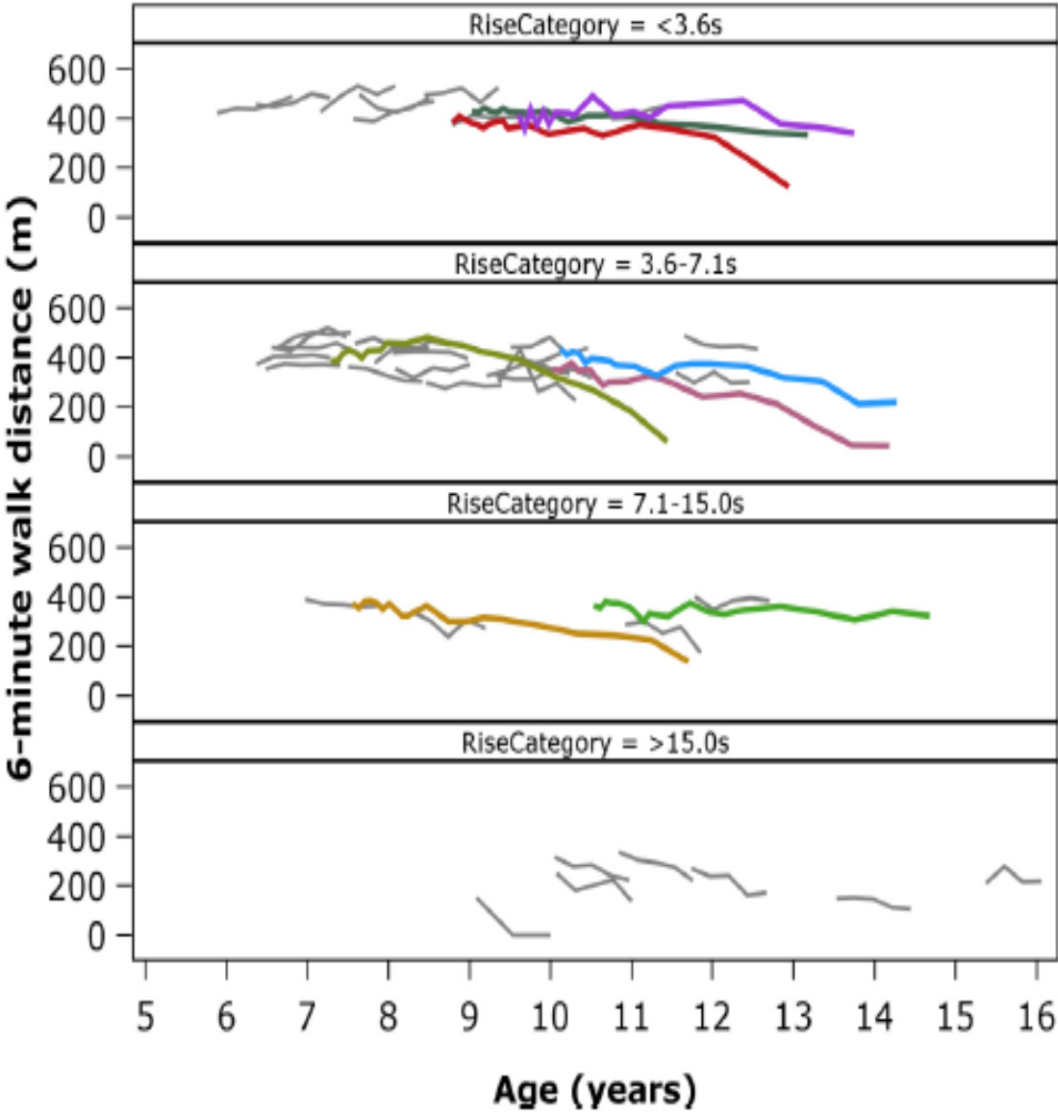


Figure 9: 6MWT, Prednisone-treated patients

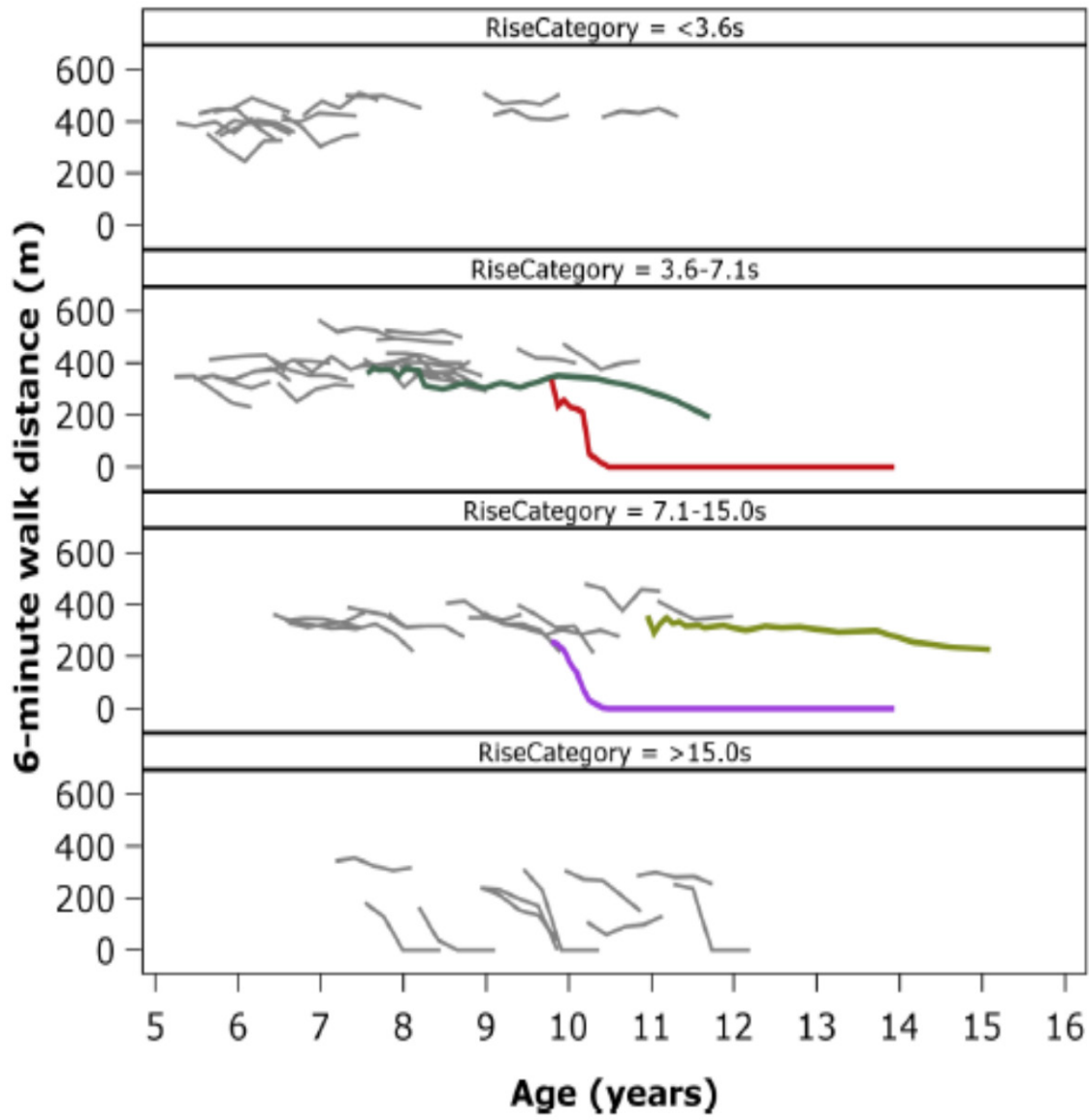
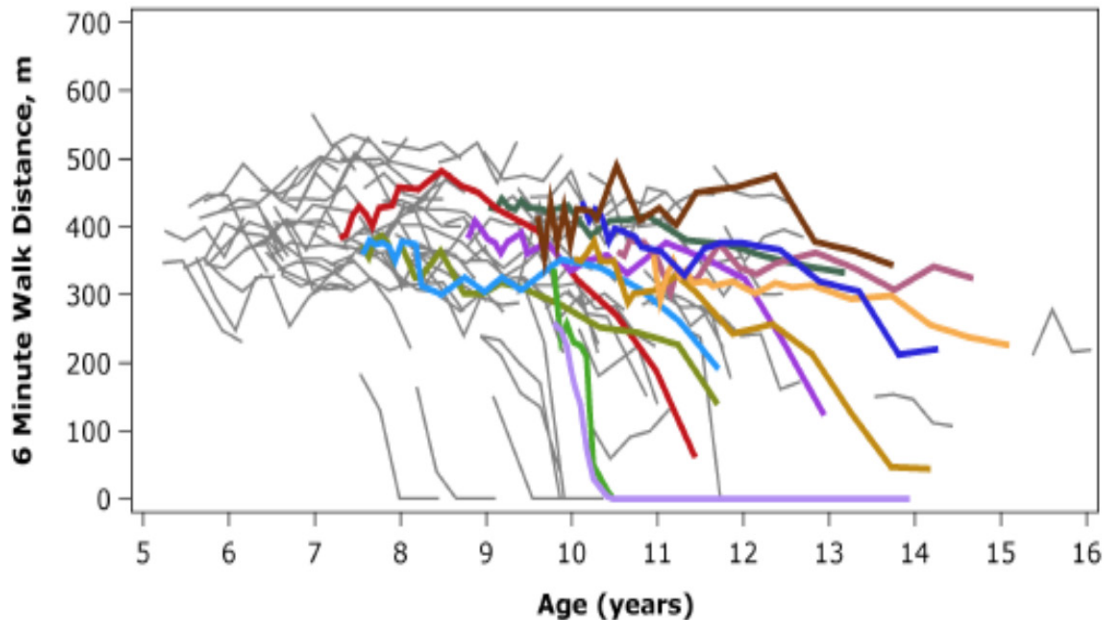


Figure 10: 6MWT, eteplirsen vs controls on placebo, all patients



Note: Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statements, “Arguably, placebo-treated patients who were blinded to treatment assignment from other controlled trials are more appropriate as matched controls than registry patients, as they may receive special care and attention as trial participants, and may be more highly motivated” as a key inaccuracy.

Sarepta: “The placebo patients from another study as referenced by the FDA are not appropriate for comparison with the eteplirsen-treated patients:

Baseline characteristics are not comparable between eteplirsen and the proposed placebo group:

- Placebo group included boys <7 years old
- Placebo group included many patients with baseline 6MWT >440 meters which is outside the eteplirsen trial’s inclusion criteria.”

FDA response: The FDA figures match patients with comparable baseline characteristics to eteplirsen-treated patients. Control patients with similar baseline characteristics to eteplirsen patients can be readily identified by examining the figures, as can the control

patients who do not match the eteplirsen patients, for example those who are younger or had a baseline 6MWT >440 meters.

Sarepta: “By virtue of the ambulatory requirement at study entry, older placebo patients (e.g., ≥11 years) were a group of pre-selected, better performing subjects”

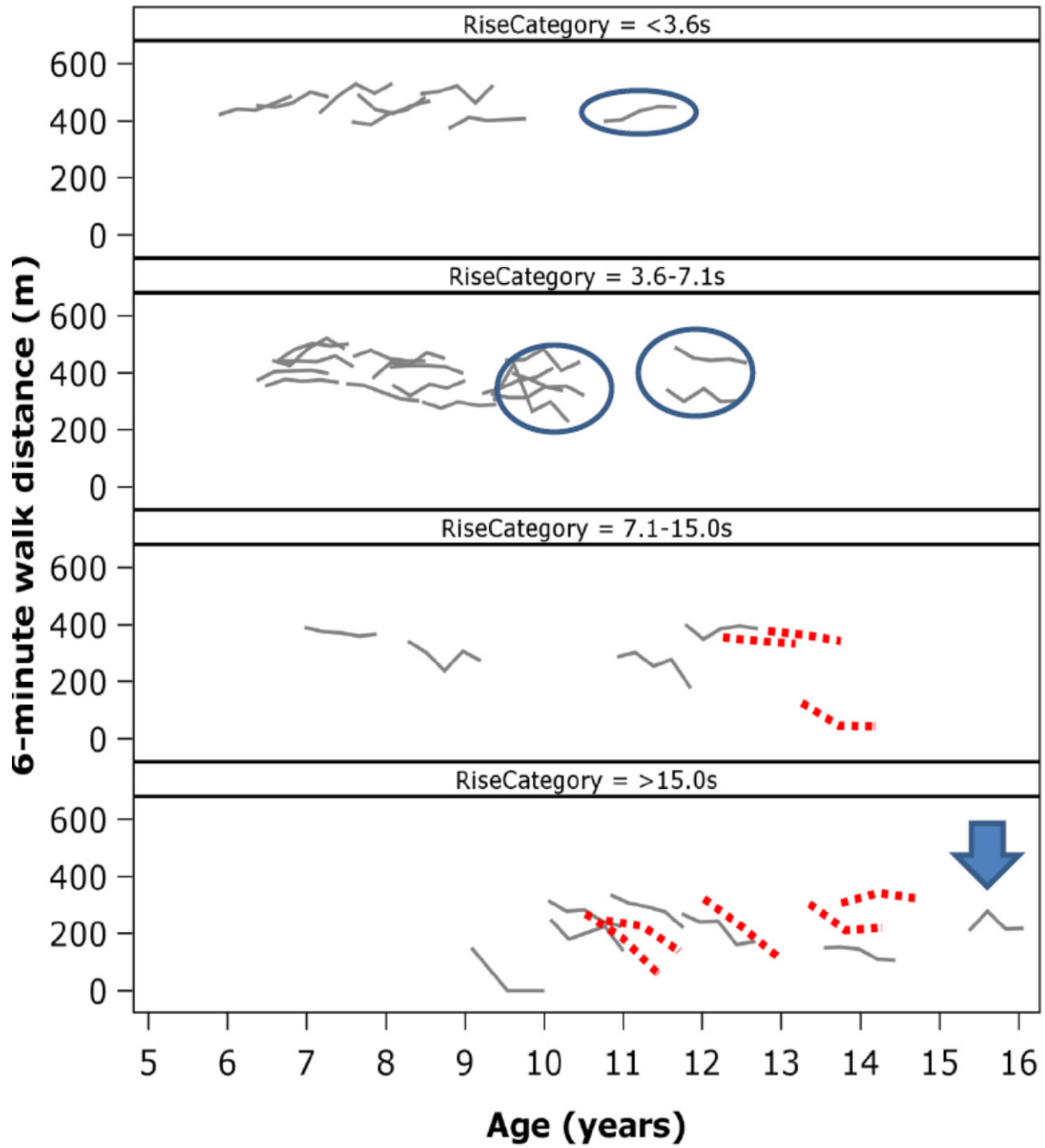
FDA response: The drisapersen placebo control patients are informative of the variability and range of function in exon-51 skippable patients. A key observation is that exon 51-skippable patients can maintain ambulation, and experience a relatively slow decline in ambulation, through an older age than is sometimes recognized.

Sarepta: “The first year of an 11-year-old-at-baseline placebo patient (i.e., 11-12 years old) to the third year of a 9-year-old boy with 3 years of eteplirsen treatment (i.e., 11-12 years old) is not a valid comparison due to the difference in duration of observation, as well as the biased selection of the 11-year-old ambulatory placebo by, irrespective of both patients having the same age at last assessment”

FDA Response: FDA did not make this comparison. The drisapersen control patients can be used to show the presence of exon-51 skippable patients who are similar to eteplirsen-treated patients. The earlier version of this memo explained that most eteplirsen patients are currently in the >15 s rise time category and can be compared to the >15 s rise time group of control patients. This comparison is now explicitly shown in Figure 11, which overlays the third year of data from eteplirsen patients with placebo patients matched on the basis of rise time at the beginning of the third year of treatment (for clarity, only deflazacort-treated patients are shown). The following are some notable observations:

- Many placebo patients in the highest (worst) rise time category show a relatively slow decline in ambulation similar to that seen in many of the eteplirsen patients in their third year of treatment, including placebo-treated patients who are as old or older than the eteplirsen-treated patients (e.g., Figure 11, arrow).
- Increase in rise time generally occurs prior to loss of ambulation. Many placebo patients in lower (less advanced) rise time categories would be predicted to maintain ambulation for several years (Figure 11, circles).

Figure 11: Third-Year Eteplirsen 6MWT (Deflazacort-treated patients)



Sarepta: “Comparison of eteplirsen-treated patients to the appropriately matched external control shows that more than one year is required to observe a divergence in disease progression between the two groups”

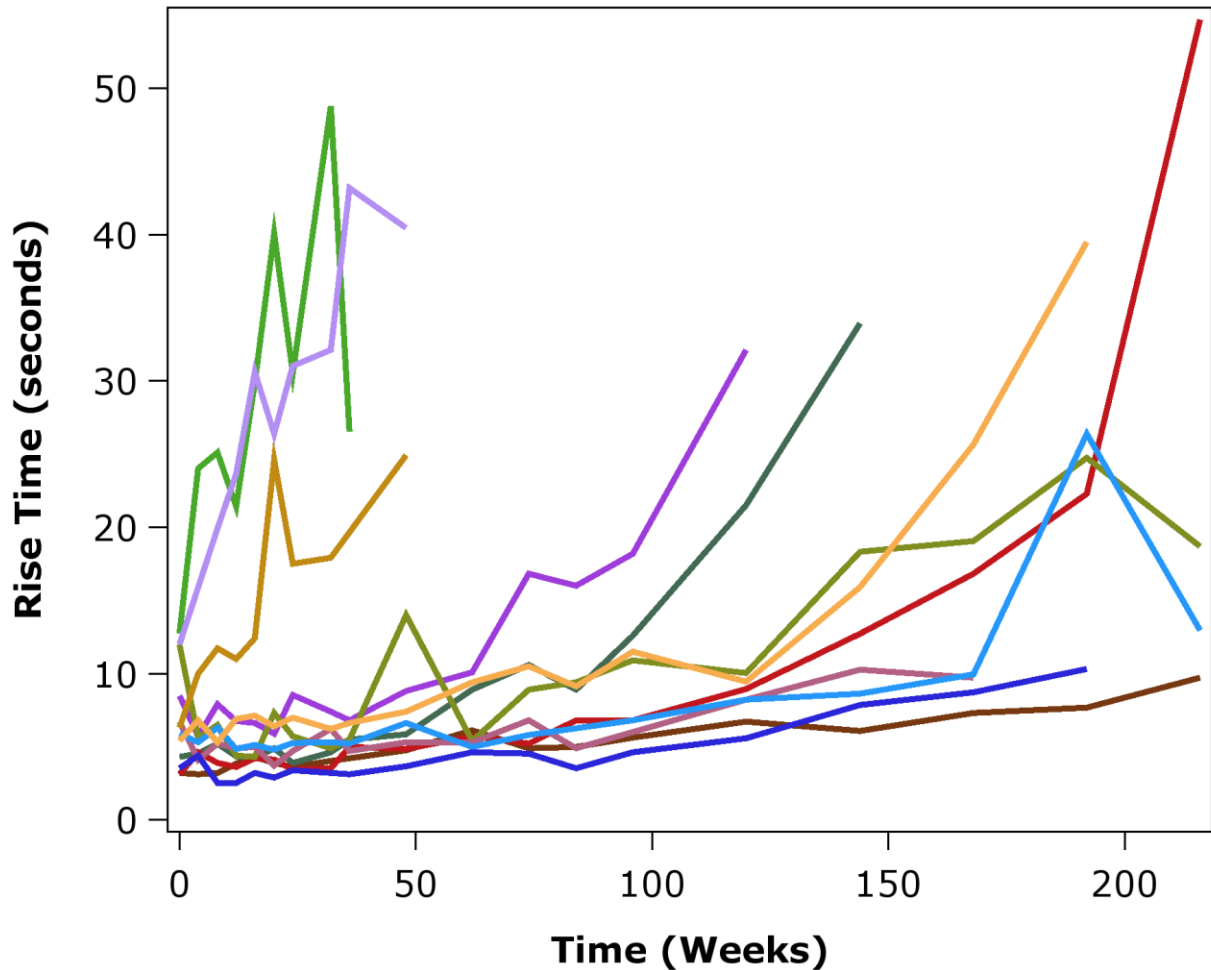
FDA response: The comparison to placebo controls incorporates the full duration of eteplirsen treatment and all potential cumulative effects. After 3+ years of treatment, eteplirsen patients are still within the range of clinical condition that occurs in the natural history of exon-51 DMD.

Because evidence that even a few eteplirsen patients might have progressed markedly differently than expected by natural history would be of interest, a few additional observations about these data are important. Assignment of eteplirsen patients to rise-time category is affected by random noise in the baseline measure. Specific patients may appear to progress faster or slower than “matched” controls, but the noise inherent in matching needs to be considered. For example, the patient indicated by the bright green line in Figure 8 was placed in the 7.1- to 15-second rise time category, but had large variability for rise time values, and a more accurate estimate of rise time for this patient might be closer to 5 seconds, suggesting that matching to a less advanced group of historical controls might have been as, or more, appropriate. In addition, a number of other factors can confound efforts to match treated with historical patients. For example, the sponsor has argued that loss of muscle, as measured by MRI, was more severe at baseline in two patients than suggested by functional tests, decreasing the interpretability of the rapid loss of ambulation experienced by these patients after starting eteplirsen.

c. Increases in rise time in eteplirsen-treated patients predict a high likelihood of sequential loss of ambulation within 1 or 2 years

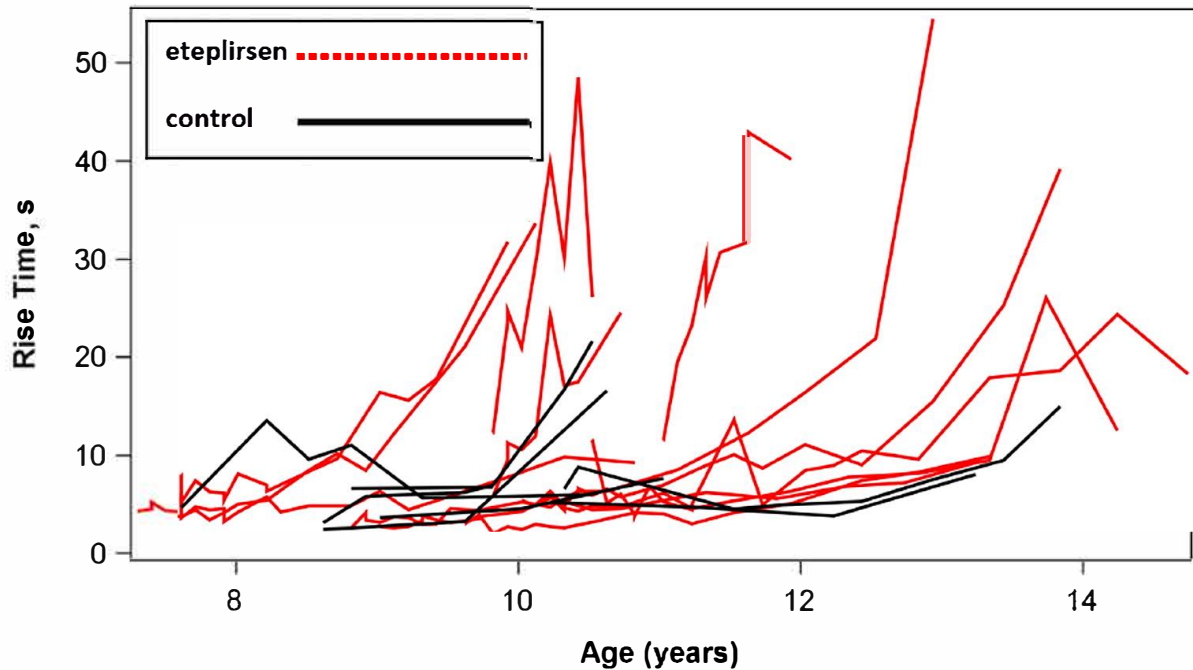
Figure 12 shows rise time from floor for the eteplirsen patients. Three eteplirsen patients lost the ability to rise from the floor in the first year of Study 201. The applicant has, at times, proposed that after an initial time period in which dystrophin levels from eteplirsen accumulated, disease progression largely stabilized in treated patients. All patients in Study 202 have continued to progress steadily while taking eteplirsen, as indicated by rise time from floor, without any discernible stabilization or slowing. Most have now become unable, or nearly unable, to rise from the floor, which predicts a high likelihood of sequential loss of ambulation within 1 or 2 years.

Figure 12: Rise Time, Study 201/202



Rise-time data were submitted by the applicant for 8 of their 13 natural history patients, and new FDA analyses are shown in Figure 13 for the comparison with rise time data in eteplirsen-treated patients. In the graph, a more horizontal slope indicates a slower rate of progression, whereas a faster rate of progression is indicated by a more vertical slope. Progression of rise time was marked by a high level of inter-patient variability, but was generally similar for eteplirsen and natural history patients. Note that two of the patients with the most preserved rise time were historical control patients, and that no eteplirsen treated patient declined slower (more horizontal course) than the range set by the natural history patients.

Figure 13: Rise Time, Eteplirsen in Study 201/202 and External Controls

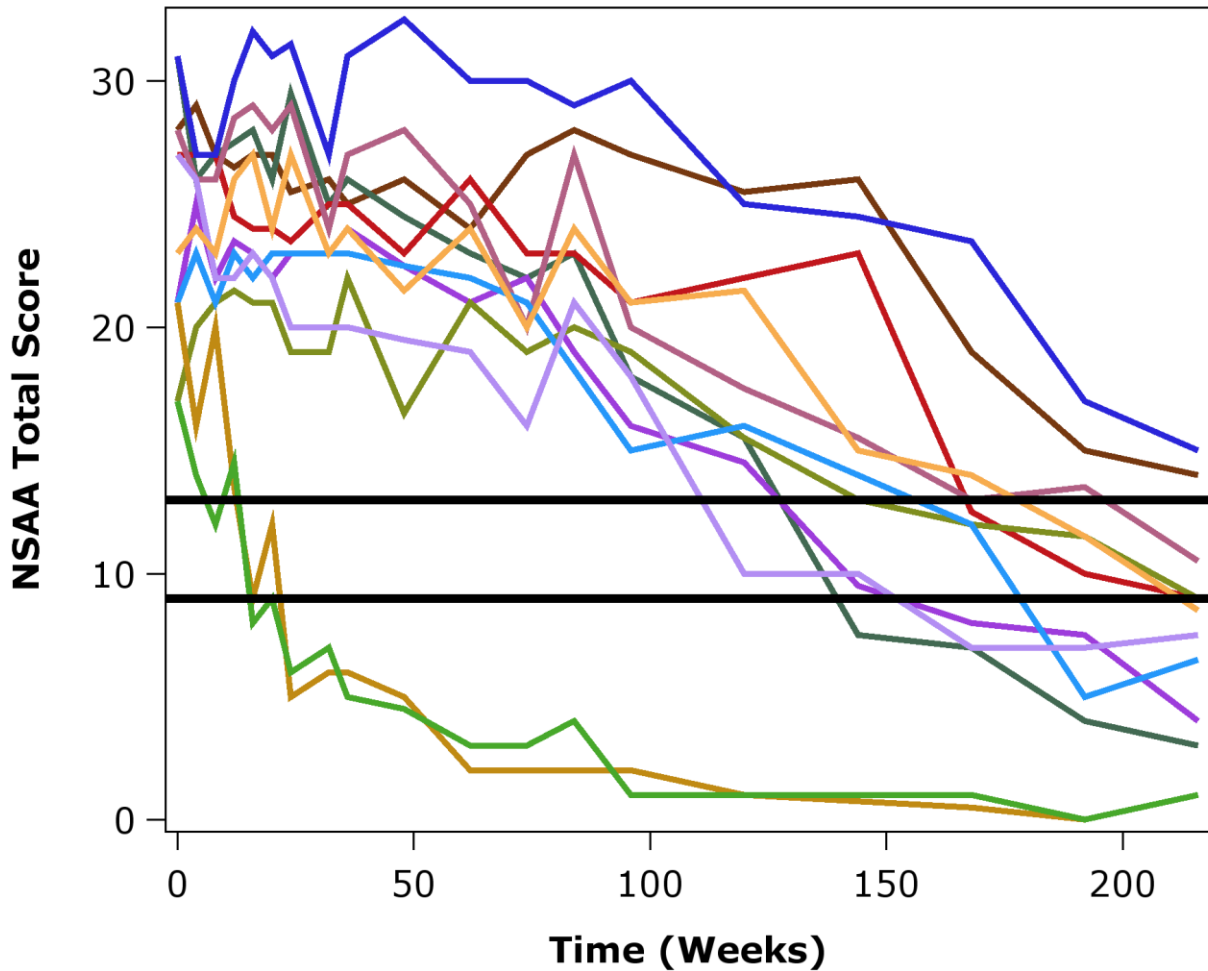


The applicant has emphasized a time-to-loss analysis for rise time but, similar to 6MWT, the recording of when a function is lost is partly subjective, and may be substantially affected by the level of disability at which the examiner concludes that attempting the test of function is no longer warranted. The data in Figure 13 suggest that rise time may have been measured through a higher degree of disability for eteplirsen-treated patients, through rise times into the 40- and 50-second range, whereas above a rise time of about 20 to 25 seconds, control patients may have been considered unable to perform the task by the examiner.

Similar observations [steady progression] were noted for NSAA, which measures broader abilities related to muscle strength that are important for walking, including standing from a chair and ability to climb on and off a box step. As NSAA score decreases, patients may still be able to walk, but are at greater risk of falls, less able to assume a safe position if a fall occurs, and less able to stand up after falling. Eteplirsen patients declined by roughly 5 points/year on average (Figure 14), similar to patients in the NorthStar network. The two horizontal lines in Figure 14 indicate NSAA scores of 9 and 13 that have been reported to be associated with being

either 1 or 2 years, respectively, from loss of ambulation.⁴⁵ Combined with loss of ability to rise from the floor, the NSAA scores suggest that the eteplirsen patients, who are currently 11 to 14 years or age, are at, or close to, a level of muscle strength often associated with use of a wheelchair.

Figure 14: NSAA, Study 201/202



d. Issues with comparison of eteplirsen-treated patients with applicant’s proposed historical controls

Untreated historical control groups tend to have worse outcomes than apparently similar control groups in randomized studies. Patients in randomized studies need to meet certain

⁴⁵ Ricottii et al (2015) The NorthStar Ambulatory Assessment in Duchenne muscular dystrophy: considerations for the design of clinical trials. J. Neurol Neurosurg Psychiatry. 0,1-7.

criteria to be entered that generally select a less sick population than is typical of external control groups. Such concerns appear to apply to muscular dystrophy, although the magnitude of this effect is difficult to quantify. In patients with fascioscapulohumeral muscular dystrophy Statland et al.⁴⁶ observed that “whereas natural history data showed a *decrease* in strength over 1 year, there was an apparent *increase* in strength at 6 months in 2 of the 3 clinical trials in both the placebo and treatment groups.” [emphasis added] The authors concluded that this type of bias should be taken as a reminder of the importance of placebo groups when measuring strength in muscular dystrophy.

Supportive care can prolong ambulation in DMD by several years, but its effectiveness is dependent on both type and intensity of care, which is likely to differ substantially between patients enrolled in observational studies or registries versus interventional treatment studies. DMD care guidelines specify that corticosteroid efficacy needs to be balanced with side effects in the context of the individual patient’s goals. Patients enrolled in efficacy trials would likely be more interested in maximizing steroid efficacy compared to patients enrolled in observational natural history studies. This appears to have been the case for the eteplirsen patients compared to the controls selected by the applicant. A higher proportion, 69% vs. 8%, of the natural history controls vs. eteplirsen patients were on regimens other than daily dosing that are often selected to decrease side effects but that are thought to be associated with less efficacy. Doses of corticosteroids also appear to have been lower in the applicant’s natural history patients, which included those “in whom the dose had not been always completely adjusted to the current weight.”⁴⁷ Adherence to treatment guidelines is difficult to measure, but adherence in the eteplirsen study was reported to be exceptional, while there is evidence that care received in the regions of origin of many of the sponsor’s historical control patients was likely of lower intensity.⁴⁸ Finally, as the sponsor’s natural history study proceeded, some patients left to enter interventional clinical trials, further decreasing the similarity of the natural history cohort to the eteplirsen patients.

⁴⁶ Statland JM et al (2013) Reevaluating measures of disease progression in facioscapulohumeral muscular dystrophy. *Neuromuscul Disord.* 23,306-12.

⁴⁷ Mazzone E et al (2010) North Star ambulatory assessment, 6-minute walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscular Disorders.* 20,712-716.

⁴⁸ Landfeldt E et al (2015) Compliance to care guidelines for Duchenne muscular dystrophy. *Journal of Neuromuscular Diseases.* 2,63-72.

Note: Following public release of the original FDA briefing document on 1/15/16, the applicant provided an addendum to their briefing document that highlighted the statements, “Finally, as the natural history studies proceeded, some patients left to enter interventional clinical trials, further decreasing the similarity of the natural history cohort to the eteplirsen patients” as a key inaccuracy.

Sarepta: “Two types of missing data sensitivity analyses were performed, the results confirmed that the magnitude of difference remained over 100 meters and nominal statistical significance was maintained:

- MMRM using all the available data
- Last Observation Carried Forward imputation (conservative analysis assuming that the 2 control patients did not decline)”

FDA response: The applicant’s response, describing two types of analyses used to impute missing data, suggests that they construed FDA’s concern to be the problem of missing data, i.e., missing data from patients who left the natural history study. But FDA did not make this point to highlight missing data as an issue. FDA’s intent was to underscore the inherent and profound difference between patients in the interventional eteplirsen trial and patients in the observational study.

There are many reasons to conclude that there were meaningful differences between the groups, both at baseline and during the conduct of the study. Some additional examples of specific concerns are listed below.

- Important aspects of supportive care were incompletely and/or incorrectly recorded for both Study 201/202 patients and historical controls:

After FDA noted there were potentially clinically meaningful differences in steroid treatment between eteplirsen treated and control patients, the applicant revised the raw data for historical control patients, stating that it was incorrect and/or incomplete as originally submitted to FDA: one patient was changed from “intermittent” to “continuous” treatment, and 3 were changed from “unknown” to “continuous.” The reliability of data revised in this way is questionable. In the setting of knowledge of treatment arm, changing source data can introduce bias in favor of drug-treated patients. Applicants may be more likely to selectively question and revise data to support the apparent drug effect. For example, FDA recently received from the applicant source documents containing data

on steroid use by the natural history patients in Belgium, indicating that one patient was initiated on only 6 mg/day deflazacort, apparently due to a misunderstanding, but this was not brought to FDA's attention.

There remains reason to be concerned that the differences in steroid treatment may have impacted prognosis. For example, steroids were reported to have been initiated in eteplirsen treated patients at a younger age than for historical controls (on average, over one year earlier). The possible impact of that difference on clinical outcomes is impossible to assess, which again highlights the limitations of the comparison to historical controls.

- *Supportive care was not well documented for the eteplirsen-treated patients in Study 201/202. In response to an FDA request of 20 August 2015 for additional details about supportive care, the applicant responded "the study 368-us-201 and 4658-us-202 protocols did not include collection of supportive measures such as the use of night splints, physical therapy, etc., in the study population."*

Patient compliance with clinical recommendations is not expected to be complete, and there is a concern that it would be higher in interventional compared to observational studies. In the limited source documentation available for the historical control patients, some difficulty gaining patient compliance is documented.

- *In a recently published correction,⁴⁹ the investigators of the Italian natural history study that contributed 10 of 13 historical control patients reported substantial changes in accounting for basic aspects of the patient registry – e.g., patient numbers, duration of enrollment, dropouts, survival, etc. Such changes raise concern about the reliability of the data, and that efforts to correct the data may have been influenced by investigator expectations about the disease course. In addition, the revised numbers indicate a high percentage of assessments were not carried out at 36 months (about 40%), increasing concern that the data collected might not have been representative of the original population.*

The original and corrected statements are as follows [emphasis added]:

⁴⁹ Pane et al. (2015) Correction: long term natural history data in ambulant boys with Duchenne muscular dystrophy: 36-month changes. PLOS ONE 10(12):e0144079.doi:10.1371/journal.pone.0144079

- *ORIGINAL: Of 113 patients who fulfilled the inclusion criteria and entered the study, **96** also had an assessment at 36 months. One died, **2 were lost at follow up** and the other **14** entered interventional clinical trials*
- *CORRECTED: Of 113 patients who fulfilled the inclusion criteria and entered the study, **70** also had an assessment at 36 months and another 26 were new patients, enrolled with the same criteria. Of the 43 patients excluded from the second year, 17 had not reached the 3 year assessment, 4 had assessments at different times but not at 3 years because they entered natural history clinical studies, 5 were younger than 5 years at baseline, **9 were lost at follow up** and **8** entered into a clinical study*
- *Study protocols for the Italian and Belgian observational DMD registries were brief and lacked detail, including the criteria by which it would be determined whether a patient should be deemed unable to complete an endpoint measure without attempting the test.*

Recent evidence from the Cooperative International Neuromuscular Research Group (CINRG) reinforces the observation that seemingly small differences in steroid treatment and clinical care may have relatively large effects, up to *several years*, on age at loss of ambulation.⁵⁰ The CINRG investigators caution that “differences in standards of care and dosing complicated interpretation...this study emphasizes the necessity of a randomized blinded trial of GC [glucocorticosteroid] regimens in DMD.” This is an important conclusion for DMD drug studies more broadly because differences of several years in age of loss of ambulation among different groups of patients may not be large enough to determine reliably the contribution of a drug versus other factors.

The table below shows some of the numerical data from the CINRG study that is referred to in the paragraph above. There is a difference of about 3 years in median age of loss of ambulation between two large groups of patients, one treated with prednisone and the other with deflazacort. Also notable is that loss of ambulation differed by 2 years between patients on differing deflazacort dosing schedules, perhaps reflecting a combination of factors including random effects from small sample size (N = 8 for one group). Bello et al also note that “DFZ [deflazacort] is not commercially available in the United States, where many CINRG sites are located, and it is more expensive than

⁵⁰ Bello L et al (2015) Prednisone/prednisolone and deflazacort regimens in the CINRG Duchenne natural history study. *Neurology*. 85, 1048-1055.

prednisone, implying that its use may have been associated with higher standards of care and possibly adherence.”

Steroid/Regimen*	Median loss of ambulation (years)	N
Prednisone/Daily	11	94
Deflazacort/Daily	14	80
Deflazacort/Switched	16	8

*daily vs. weekly

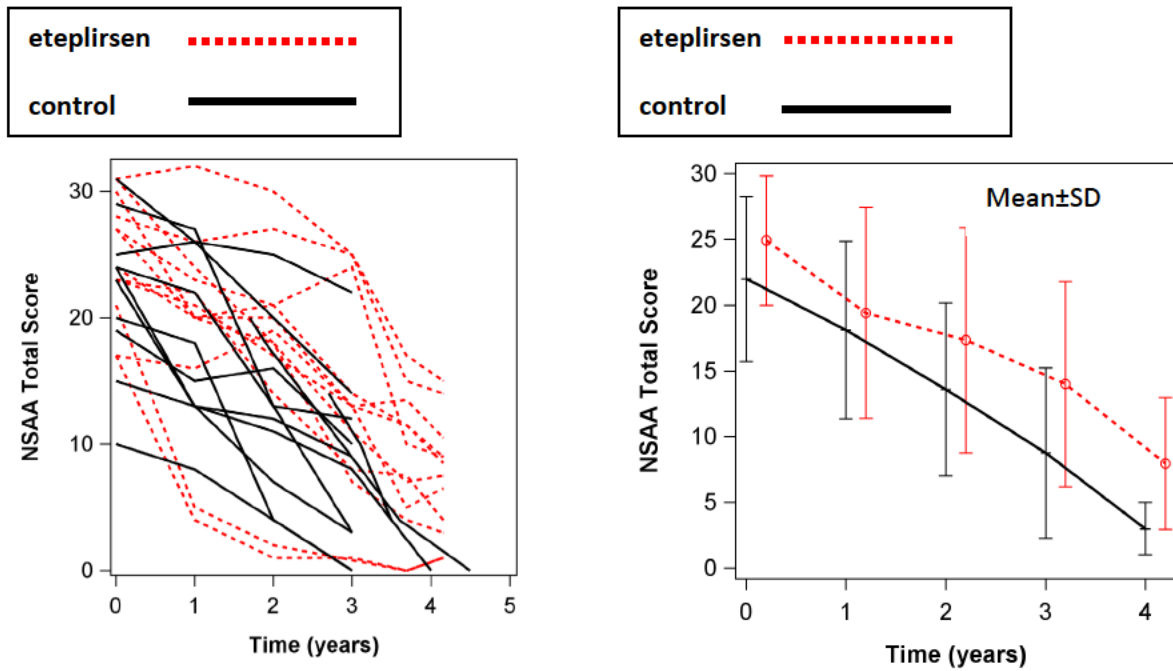
e. NSAA, Eteplirsen vs. Applicant’s Controls

Comparison of eteplirsen patients (red) to the applicant’s historical controls (black) is shown for NSAA in Figure 15 for individual patients (left) and mean for each group (right).

In source documents recently received from the applicant, there appears to be documentation that NSAA was, in a number of instances, recorded as zero for the applicant’s historical control patients without being measured, potentially underestimating the patient’s actual abilities. The applicant identified 2 instances, and initial FDA review suggests there may have been more.

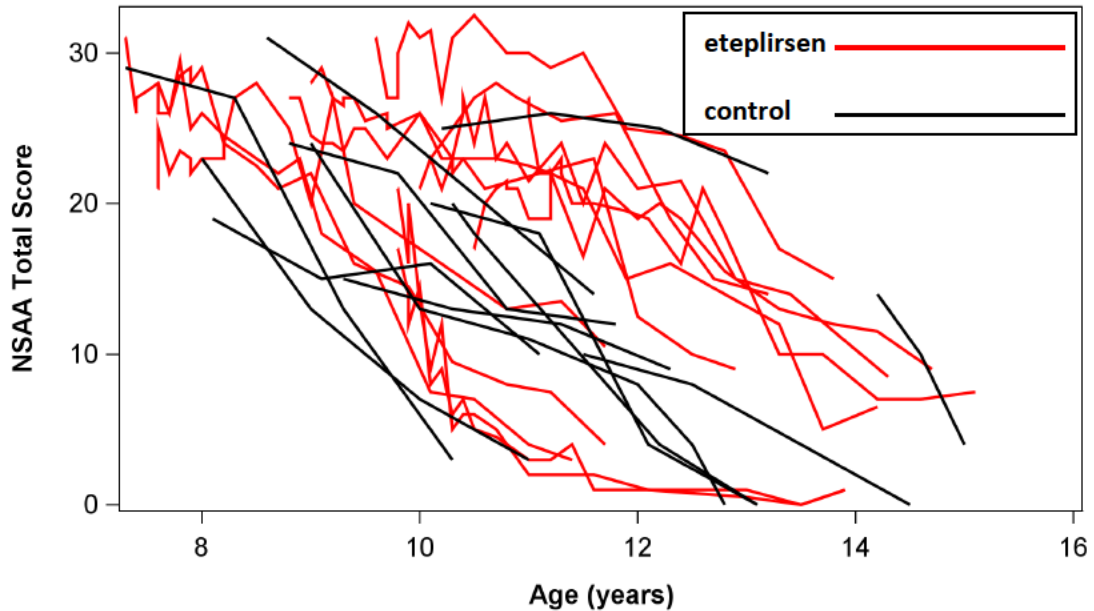
As discussed above, the effects of bias can be considerable in historically-controlled trials, with many factors potentially favoring the treatment arm. The similarity of the clinical course of patients is therefore notable. The similarity between the groups on NSAA and, in particular, the large magnitude of the standard deviations, suggest that eteplirsen does not have the type of large beneficial effect that would be possible to reliably detect in even a well-designed historically-controlled trial.

Figure 15: NSAA, eteplirsen vs applicant's historical controls



Because muscle function is strongly correlated with age in DMD, Figure 16 displays NSAA vs. age (in contrast to vs. years on treatment) to provide a better matched comparison of patients. NSAA values for control patients occur over the entire range of values for eteplirsen patients, e.g., two of the patients with the most preserved NSAA score at both age 13 and 14 years are external control patients.

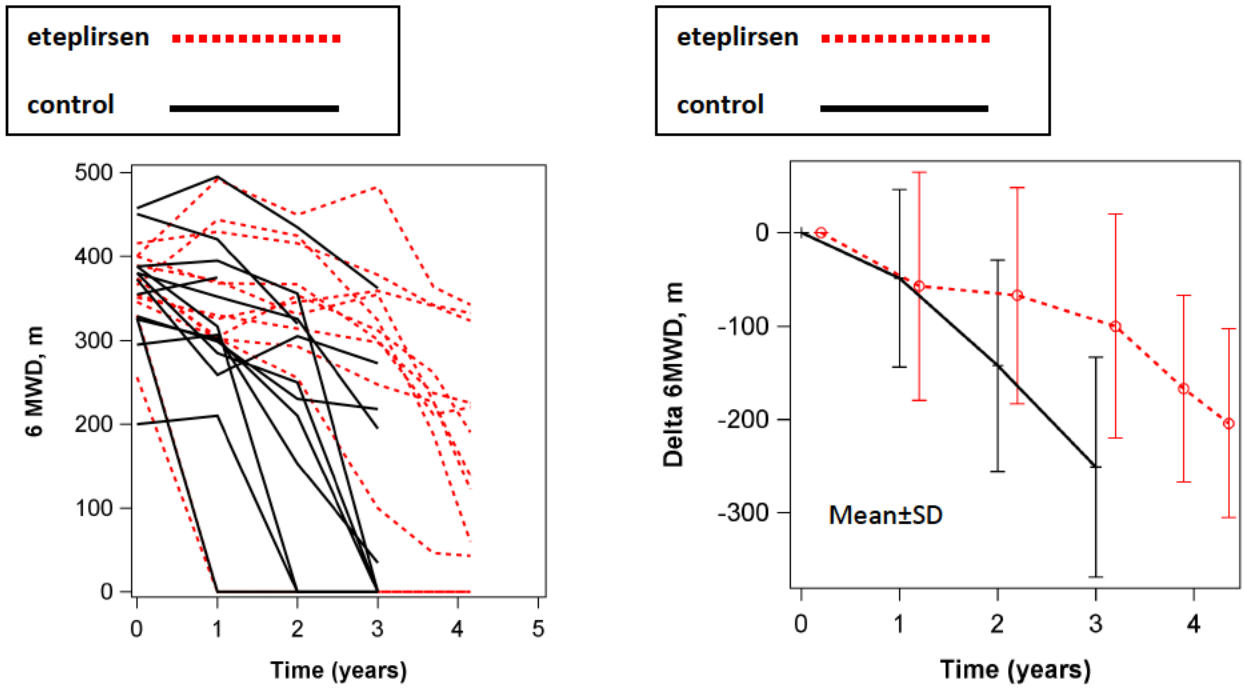
Figure 16: NSAA by Patient Age



f. 6MWT, Eteplirsen vs. Applicant's Controls

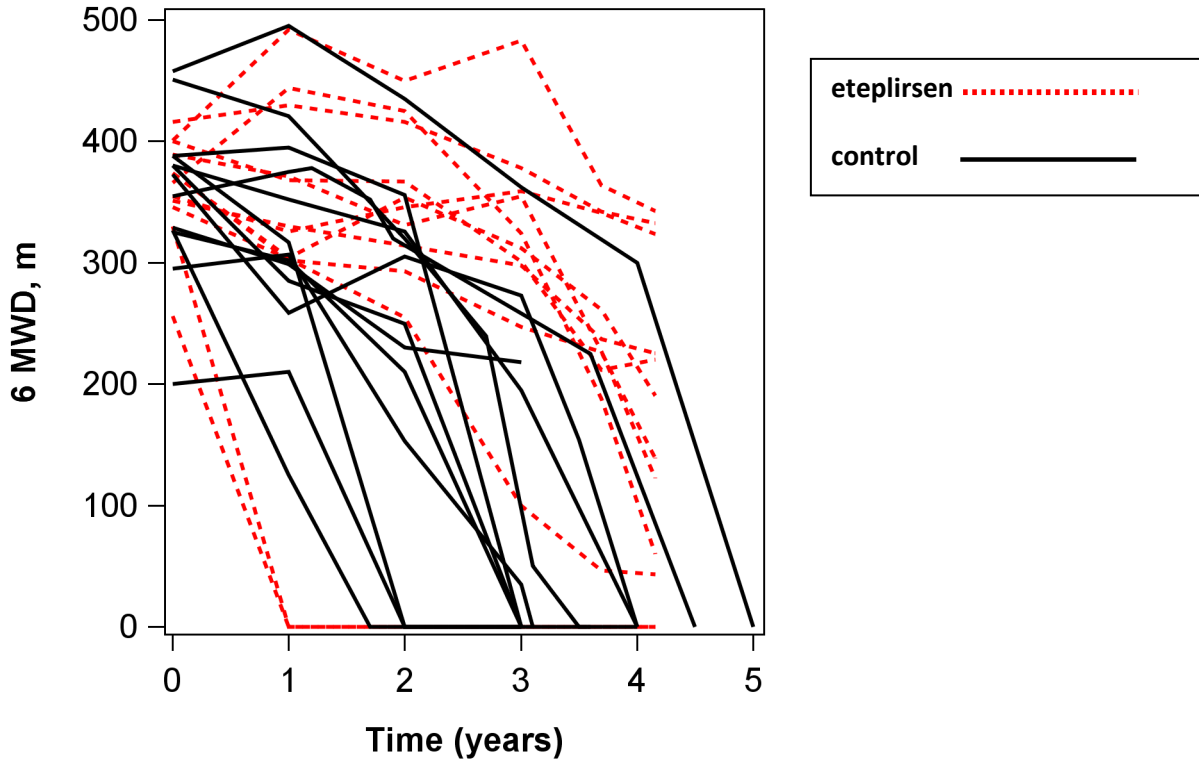
Comparison of eteplirsen patients (red) to the applicant's historical controls (black) is shown for 6-minute walk distance in Figure 17, for all patients (left) and mean for each group (right). As discussed above, FDA has long expressed concern to the applicant that the 6MWT is particularly susceptible to bias, and unreliable in Study 202. Importantly, whereas the difference in 6-minute walk distance shown would be of clinical importance if observed in a double-blind, placebo-controlled trial, the finding is extremely difficult to interpret given all of the limitations of historically-controlled trials noted above.

Figure 17: 6MWT, eteplirsen vs applicant's historical control



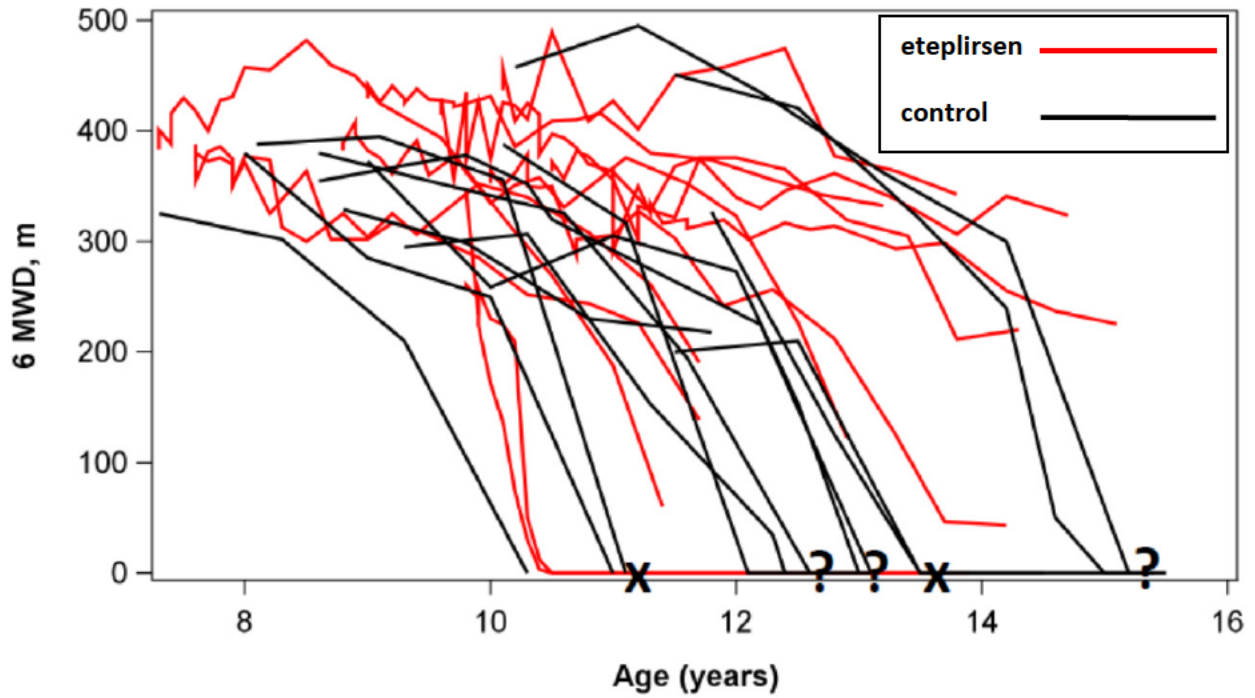
An updated version of 6MWT vs. time on treatment/observation is shown in Figure 18.

Figure 18: 6MWD vs. Years Observed



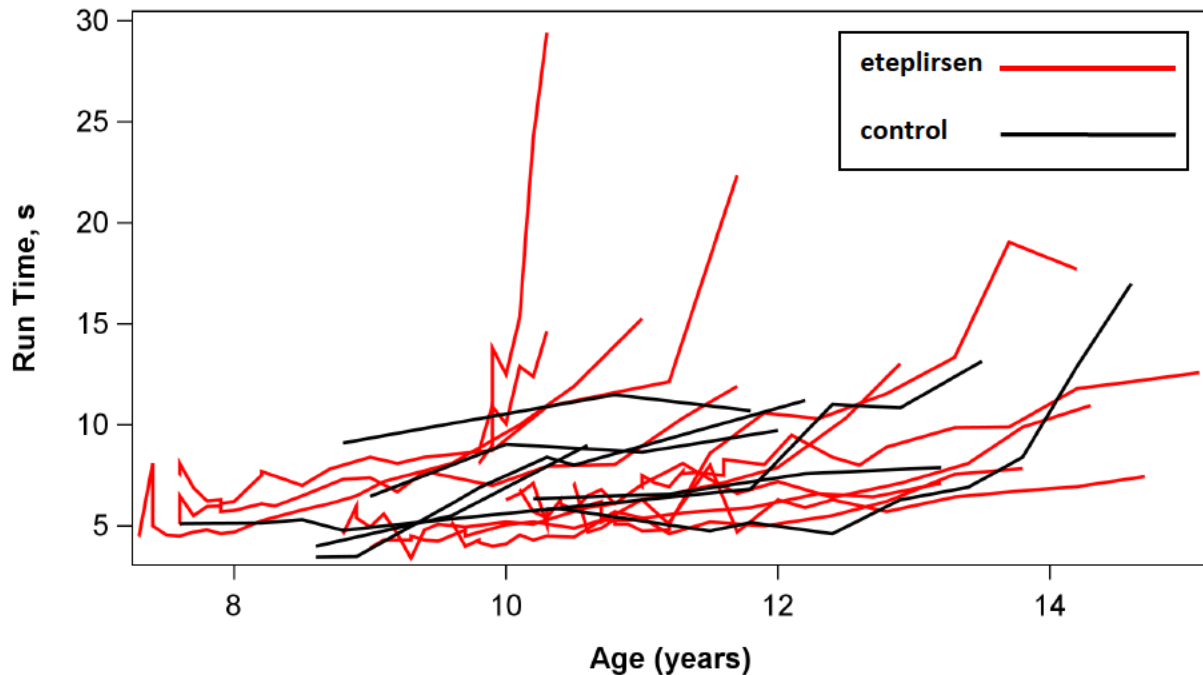
Because function is strongly correlated with age in DMD, Figure 19 displays 6MWT values vs. age (as opposed to years on treatment) to provide a better-matched comparison of patients. A majority of eteplirsen patients (red) are declining in close parallel to the paths of historical control patients of similar age (black). For the patients older than 14 years, several eteplirsen patients are ambulating at a time when control patients of similar age have 6MWT values of zero, but as noted above, a number of these values appear not to represent the true ambulatory abilities of the patients (in the figure “x” marks patients who were ambulatory but recorded as having 6MWT of zero, and “?” indicates patients who were reported, but seemingly not measured, to have 6MWT of zero).

Figure 19: 6MWT by Patient Age



Data for 10 m walk/run were submitted by the applicant for 7 of their 13 natural history patients, with new FDA analyses comparing eteplirsen and natural history patients shown in Figure 20. In the figure, a more horizontal slope indicates a slower rate of progression, whereas a more vertical slope indicates a faster rate of progression. Progression as measured by 10 m walk/run was marked by a high level of inter-patient variability, but was generally similar for eteplirsen and external control patients.

Figure 20: 10 m walk/run, eteplirsen and natural history patients



FDA recently received source documents that increase concern that 10m walk/run may have been measured differently for eteplirsen-treated patients compared to natural history patients. Eteplirsen patients appear to have been recorded for the time to **run** 10 m whereas at least some external control patients were recorded for the time specifically to **walk** 10 m, e.g., for one patient, a time of 7 s was recorded to **run** 10 m, and a time of 11 s was recorded to **walk** 10 m, with only the slower time submitted to the NDA as the 10 m run/walk time. Patients in the external control group who walked, rather than ran, for the test would tend to improve the results in the eteplirsen group relative to the external control group.

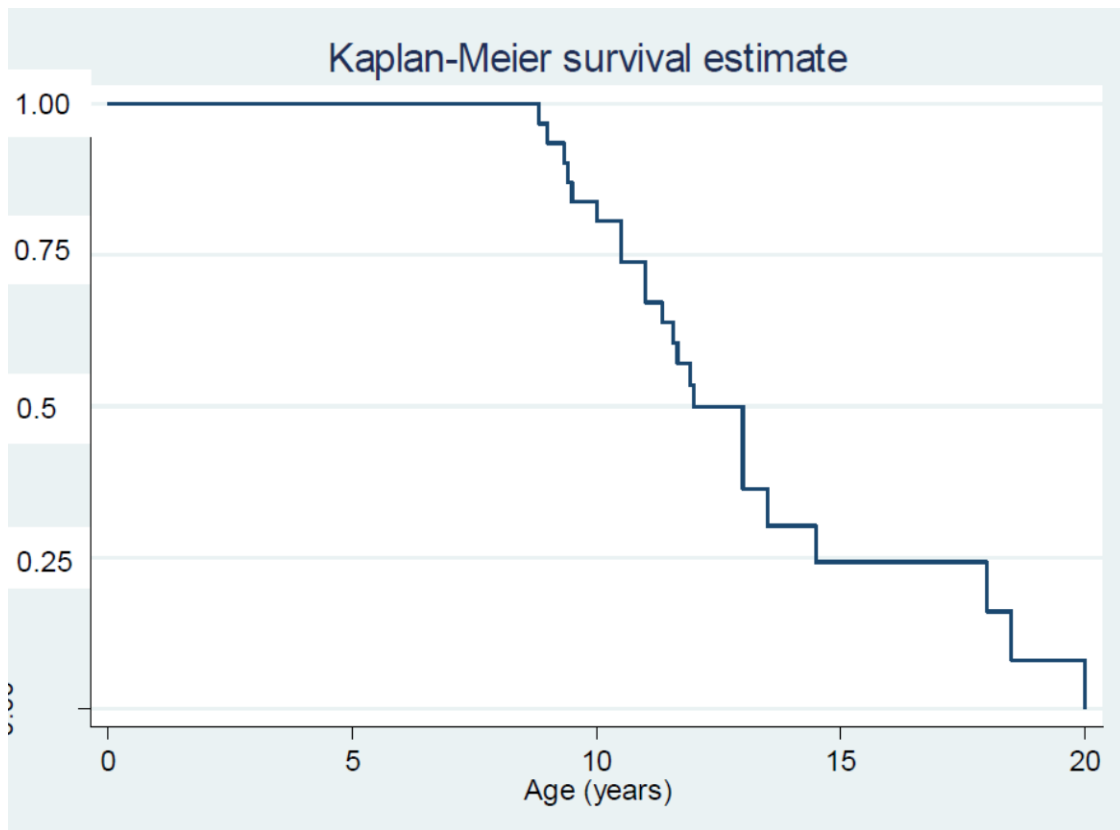
New CINRG Data: Age Range of Ambulation of Exon-51 amenable patients

As noted above, after the release of the original version of this memo, CINRG provided additional unpublished analyses to FDA for age of loss of ambulation of exon-51 amenable patients (Figure 21; analysis as provided by CINRG). CINRG is additionally providing FDA with

patient-level data that should enable a more detailed comparison with eteplirsen treated patients.

Based on the CINRG data, about 25% of exon-51 skippable patients maintain ambulation to age 16, and about 15% of patients to age 18. The oldest eteplirsen-treated patients are currently about 15 years old, such that it cannot be concluded that the ambulatory function of any eteplirsen-treated patient exceeds the expected range of natural history. This is an important point because some of the applicant's analyses give the impression that some, or most, of the eteplirsen patients have maintained ambulation longer than could have been expected compared to natural history.

Figure 21: Age at Loss of Ambulation, Exon 51 Skippable Patients



Importantly, any comparison of the eteplirsen data to the CINRG data needs to account for the fact that eteplirsen patients, upon enrolment in Study 201, had to meet criteria based on a specific level of ambulation at an age at which some patients would have already declined to a point where they would not have met these criteria. The eteplirsen patients, therefore, represent a population enriched for patients with a better prognosis than the overall exon-51

skippable population. Therefore, the percentage of eteplirsen-treated patients who would be expected to maintain ambulation would be higher than 25% at age 16 years and 15% at age 18 years, even before considering other potential sources of difference between the groups.

Pulmonary Function

Figure 22 shows the comparison of percent predicted forced vital capacity (%FVC) in eteplirsen-treated patients (colored lines) with patients on placebo (grey lines) in controlled trials of another drug investigated in exon-51 skippable DMD patients. The course of both groups of patients is generally similar, marked by general stability or slow decline, as expected in steroid-treated DMD patients in this age range.^{51,52}

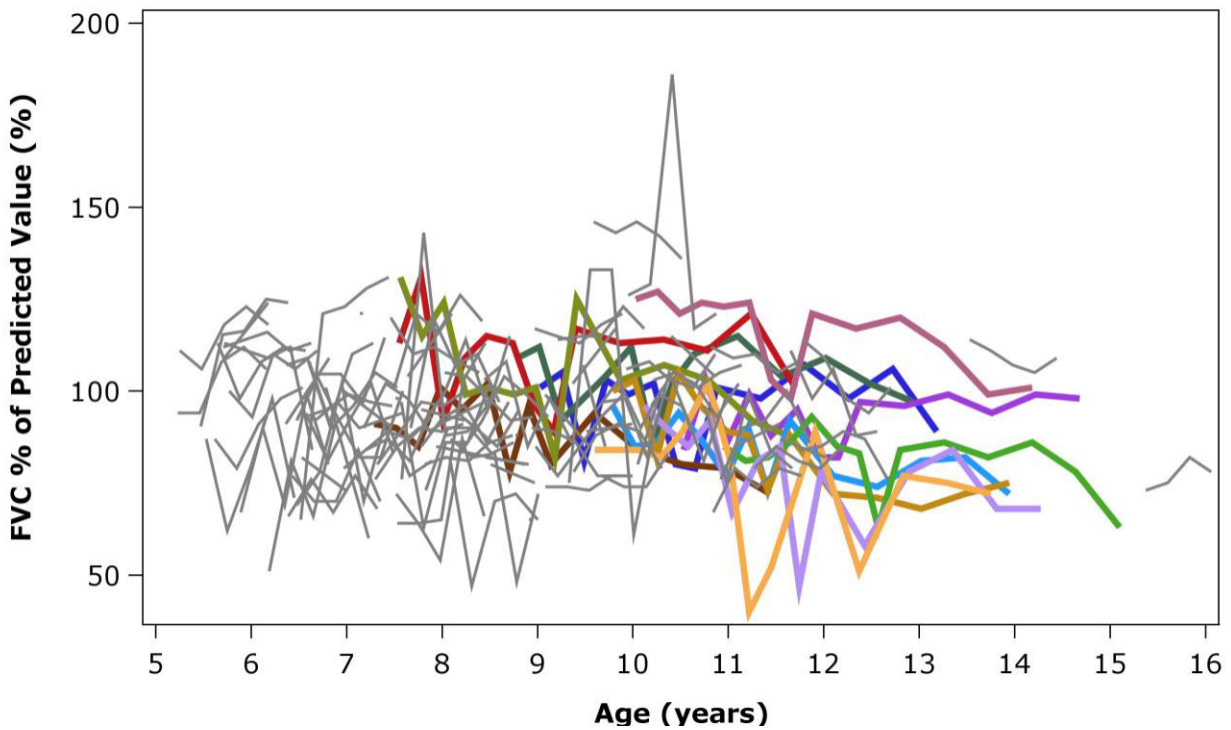
The applicant compares eteplirsen-treated patients to natural history patients who were either not treated with steroids, or who were treated for shorter periods of time. The applicant suggests steroids have little or no effect on pulmonary function, but this does not appear to be supportable.⁵³ The applicant's analyses regarding pulmonary function therefore appear to be confounded and uninterpretable.

⁵¹ Biggar WD et al (2001) Deflazacort treatment of Duchenne muscular dystrophy. Journal of Pediatrics. 138, 45-50.

⁵² Machado DL et al (2012) Lung function monitoring in patients with Duchenne muscular dystrophy on steroid therapy. BMC Research Notes. 5,435

⁵³ Gloss D et al (2016) Practice guideline update summary: corticosteroid treatment of Duchenne muscular dystrophy: report of the guideline development subcommittee of the American Academy of Neurology. Neurology 86:465-72.

Figure 22: Forced vital capacity, eteplirsen-treated patients (colored lines) vs patients on placebo in other controlled trials in exon-51 skippable DMD patients (grey lines)



g. FDA Review Team Preliminary Conclusions, Clinical Endpoints

In the context of the above, the major preliminary conclusions of the primary review team with regard to clinical endpoints are listed below:

1. The natural history of DMD in patients amenable to exon 51 skipping has been characterized in a number of observational natural history studies and controlled trials, and the range of age at loss of ambulation is very wide, currently between about 8 and 16 years for most patients. Eteplirsen patients have experienced a sequential loss of ambulatory abilities and increasing muscle weakness, as measured by rise time from floor, NSAA, 6MWT, and other tests. In the context of this considerable variability among patients, the clinical course of eteplirsen patients over more than 3 ½ years of treatment with eteplirsen has been generally similar to expected natural history of patients provided with intensive supportive care.

As noted above, recently available CINRG data suggest exon-51 skippable patients commonly maintain ambulation to older ages than is often realized, to 18 years or perhaps even older.

2. There are important differences between patients enrolled in observational natural history studies and patients enrolled in interventional drug efficacy studies, some of which are quantifiable, and some of which are not. Corticosteroid therapy appears to have been more intensive in eteplirsen patients compared to the natural history patients selected by the applicant, and this, itself, may have been capable of affecting performance. Near the time when patients lose ambulation, decisions are made by patients and caregivers about whether weakness has progressed to the point that it is in the patient's best interest to use a wheelchair to avoid the risk of falls and injuries and to decrease the effort and time required for mobility. Differences in individual care decisions, therefore, seemingly could produce large differences in 6MWT and time to loss of ambulation between eteplirsen patients and natural history controls. NSAA results, potentially representing a more direct measure of strength, suggest that differences in DMD progression between eteplirsen patients and the applicant's natural history controls were too small and variable, in the context of a poorly-controlled trial, to be reliably attributed to drug treatment.

New data and analyses described in this updated memo increase concerns about the reliability, completeness, and comparability of the clinical data for eteplirsen-treated patients and external controls. For example, differences in the way that key endpoints were measured, including the apparently large role of judgments of study personnel about when patients were deemed unable to perform an endpoint, may have underestimated the abilities of external controls. The applicant has emphasized newly submitted data on time to loss of ambulation and other functions, but such analyses appear to be particularly unreliable in the context of the differences between study arms.

Additional analyses of ambulatory functions such as rise time and 10m walk/run appear to suggest that, in the context of a poorly-controlled trial, the rate of DMD progression in eteplirsen-treated patients and external controls was generally similar. Assessing patient function in the context of age, which correlates strongly with function in DMD, may be more appropriate than by years of treatment/observation given the range of patient age enrolled in Study 201/202.

Natural history data emerging from the CINRG study suggest that a substantial percentage of exon-51 skippable patients maintain ambulation beyond 16 years, at least to 18 years of age. The oldest eteplirsen-treated patients are currently about 15 years old, such that it cannot be concluded that the ambulatory function of eteplirsen-treated patients, either as a group or considered individually, exceeds the expected range of natural history.

3. With regard to future efficacy studies, any beneficial effects of eteplirsen are unlikely to be large enough to be detectable outside of a placebo-controlled trial.

It is important to note that the exposure-response relationship of eteplirsen is not well characterized. Dose-limiting toxicity was not observed, such that higher doses of eteplirsen, with potentially greater likelihood of efficacy, could be studied in the future.

4. Clinical Safety

The clinical safety database for eteplirsen is small: 114 total patients exposed, with only 36 exposed for ≥ 24 weeks and 12 exposed for ≥ 1 year. Most of these exposures were outside of placebo-controlled studies, limiting ability to determine if adverse events were the result of drug effect or chance. However, the serious and severe adverse events that occurred were generally consistent with events expected in DMD. The 12 patients in Study 202 were exposed for >3 years, which provides some reassurance against delayed toxicity.

It is important to note that dose-limiting toxicity was not observed, such that higher doses, with potentially greater likelihood of efficacy, could be studied in the future.

In animal studies, the primary target organ was the kidney, with dose-dependent renal tubular cytoplasmic basophilia and/or vacuolation and, at the high dose, tubular degeneration/necrosis. In a mouse study, dilatation of the lateral ventricles of the brain was observed at the mid and high doses. The mechanism of this effect, and its relevance to humans, is unknown. Mean eteplirsen plasma exposures (AUC) at the NOAELs for monkey and juvenile rat were 20-fold and 6-fold, respectively, greater than that in patients dosed once weekly with 30 mg/kg IV eteplirsen.

Appendix: Applicant’s table of “Key Inaccuracies in the FDA Briefing Document”

Note: The first issue listed by the applicant in the table titled “Potential Clinical Impact” regards text from the memo from the Division and Office, and is addressed in that revised memo.

Dystrophin Analytical Methodology:

FDA Statement	Sarepta Clarification
<p><i>“It is important to note that the applicant digitally processed dystrophin images in their background material (images in Appendix 12) in such a way that low intensity values were preferentially increased to produce a higher intensity and higher contrast image.”</i> (FDA BD page 29 of PDF)</p>	<p>The digitally processed images referenced by FDA in this statement were included in Sarepta’s briefing document for demonstration purposes only, and it is far more important to note that the referenced images were not used in the analysis of fiber intensity, nor to score dystrophin-positive fibers.</p>
<p><i>“Biomarker studies on the 4th biopsy obtained at Week 180 were conducted by the applicant with technical advice from FDA. However, the reliability of results remains questionable for a number of reasons, including the lack of independent confirmation.”</i> (FDA BD page 30 of PDF)</p>	<p>Methodology for dystrophin analyses of the fourth biopsy tissue samples, including confirmatory assessments of percent dystrophin-positive fibers (PDPF) analysis performed by 3 independent pathologists, were agreed with FDA prior to conducting any analyses of the fourth biopsy tissue samples.</p> <p>In accordance with the mutually agreed-upon protocols for the assessment of dystrophin-positive fibers in DMD muscle biopsy samples from the fourth biopsy obtained at Week 180, 3 independent pathologists performed a blinded assessment of the randomized muscle fiber microscopy images, which independently confirmed the results obtained by the pathologist at Nationwide Children’s Hospital (NCH).</p> <p>Assessment of PDPF at NCH indicated a significant increase in PDPF score ($p < 0.001$) relative to untreated control samples. This increase in PDPF score was confirmed by the 3 independent pathologists ($p < 0.001$).</p>
<p><i>“Random measurement error can be large in comparison to the estimated amount of dystrophin.”</i> (FDA BD page 31 of PDF)</p>	<p>The random measurement error of our Western blot protocol for measurement of dystrophin levels was well below the observed difference between untreated and treated Week 180 biopsy samples.</p> <p>A rigorous validation of the Western blot method was reviewed by the FDA prior to Week 180 biopsy analysis. Validation data demonstrated a %CV of +/- 50% and a linear range ($R^2 > 0.9$) of sensitivity extending as low as 0.25% of normal.</p>

FDA Statement	Sarepta Clarification
<p><i>“There is no simple or reliable way to compare estimates of dystrophin amount derived from immunofluorescence with estimates derived from Western blot.”</i> (FDA BD page 35 PDF).</p>	<p>Correlation between dystrophin quantification by Western blot and IHC methods has been demonstrated by multiple laboratories (Taylor, 2012; Anthony, 2011; Anthony, 2014; Hathout, 2015 FDA Workshop on Measuring Dystrophin).</p>
<p><i>“In this context, the applicant selected three BMD patients as comparators for the Week 180 dystrophin studies, one of whom had low dystrophin level of about 2% of normal. However, the BMD patients selected by the applicant do not appear representative, and this patient may correspond to one of the rare BMD patients with very low dystrophin levels.”</i> (FDA BD page 34 of PDF)</p>	<p>BMD patient samples were not chosen to be representative; rather, they were selected in response to an FDA request to assess the relationship between dystrophin as measured by Western blot and immunofluorescence fiber intensity. Therefore, BMD samples were obtained that represented low, middle, and higher ranges of dystrophin expression. A comparable Western blot analysis-IHC correlation was presented by Hathout, et al. (MDA 2015 Scientific Conference poster, FDA-NIH workshop on measuring dystrophin, 2015), where BMD biopsies were chosen to represent low- and mid-level dystrophin expression. Consistently, their BMD low patient biopsy was 2% of normal.</p>

Potential Clinical Impact:

FDA Statement	Sarepta Clarification
<p><i>“With these two comparisons of eteplirsen to placebo, there was a positive finding for only the lower dose (30 mg/kg) and for just one of the two time points (the later time point). The lack of an effect with the higher dose group tends to undermine the finding in the lower dose group and the lack of even a positive trend at the earlier time point (with a higher dose) sheds doubt on the finding at a later time point.”</i> (FDA BD page 7 of PDF)</p>	<p>The study was designed to see whether dose (50 mg/kg vs. 30 mg/kg) or duration was the most important criterion to enable consistent dystrophin production.</p> <ul style="list-style-type: none"> • Duration of therapy was observed to be the critical variable when interpreting dystrophin levels. 12 weeks does not represent a clinically relevant duration of therapy (FDA BD page 26 of PDF). • Significant dystrophin levels were by measured at Week 24 for the 30 mg/kg dose, and, importantly, at Weeks 48 and 180 for both the 30 and 50 mg/kg doses by all dystrophin assay methods.
<p><i>“Arguably, placebo-treated patients who were blinded to treatment assignment from other controlled trials are more appropriate as matched controls than registry patients, as they may receive special care and attention as trial participants, and may be more highly motivated.”</i></p>	<p>The placebo patients from another study as referenced by the FDA are not appropriate for comparison with the eteplirsen-treated patients (FDA BD pages 8, 9, 40-44, and 50 of the PDF):</p> <p>Baseline characteristics are not comparable between eteplirsen and the proposed placebo group:</p> <ul style="list-style-type: none"> • Placebo group included boys <7 years old

FDA Statement	Sarepta Clarification
(FDA BD page 13 of PDF)	<ul style="list-style-type: none"> Placebo group included many patients with baseline 6MWT >440 meters which is outside the eteplirsen trial's inclusion criteria <p>Placebo patients were followed for only one year, whereas eteplirsen-treated patients were followed for 3 or more years:</p> <ul style="list-style-type: none"> By virtue of the ambulatory requirement at study entry, older placebo patients (e.g. ≥11 years) were a group of pre-selected, better performing subjects. The first year of an 11-year-old-at-baseline placebo patient (i.e. 11-12 years old) to the third year of a 9-year-old boy with 3 years of eteplirsen treatment (i.e. 11-12 years old) is not a valid comparison due to the difference in duration of observation, as well as the biased selection of the 11-year -old ambulatory placebo boy, irrespective of both patients having the same age at last assessment. Comparison of eteplirsen-treated patients to the appropriately matched external control shows that more than one year is required to observe a divergence in disease progression between the two groups.
<p><i>"The robustness of the study result is a concern since a single patient could change the results substantially."</i></p> <p>(FDA BD page 69 of PDF)</p>	<p>This statement is inaccurate. A comprehensive sensitivity analysis was performed in order to address any potential issue regarding robustness of the data. Specifically:</p> <ul style="list-style-type: none"> Two patients were removed: the best performing eteplirsen and the worst performing external control patient. Results demonstrated a robust 6MWT treatment advantage of >100 meters with nominal significance.
<p><i>"Finally, as the sponsor's natural history study proceeded, some patients left to enter interventional clinical trials, further decreasing the similarity of the natural history cohort to the eteplirsen patients."</i></p> <p>(FDA BD page 47 of PDF)</p>	<p>Two types of missing data sensitivity analyses were performed, the results confirmed that the magnitude of difference remained over 100 meters and nominal statistical significance was maintained:</p> <ul style="list-style-type: none"> MMRM using all available data Last Observation Carried Forward imputation (conservative analysis assuming that the 2 control patients did not decline)

Regulatory Feedback:

FDA Statement	Sarepta Clarification
<p><i>"As the duration of exposure in Study 202 increased, the applicant proposed comparing the clinical course of treated patients to historical controls."</i></p> <p>(FDA BD page 38 of PDF)</p>	<p>The proposal to compare with historical control patients originated from the FDA. Specifically, a requirement to compare the clinical course of treated patients in Study 202 to matched patient-level historical control data was made by the FDA at the March 2014 guidance meeting, and reiterated at the September 2014 pre-NDA meeting. Sarepta had proposed an open-label confirmatory study comparing treated patients to concurrent (not historical) untreated patients with exon deletions not amenable to skipping exon 51 (i.e. the PROMOVI study).</p>

IV. Statistical Review



U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research
Office of Translational Sciences
Office of Biostatistics

STATISTICAL REVIEW AND EVALUATION

CLINICAL STUDIES

NDA/BLA #: 206488

Drug Name: EXONDYS 51™ (eteplirsen)

Indication(s): Duchenne muscular dystrophy (DMD)

Applicant: Sarepta

Date(s): Submission date: 6/26/2015
PDUFA Date: 2/26/2016

Review Priority: Priority Review

Biometrics Division: Division I, Office of Biometrics (HFD -710)

Statistical Reviewer: Xiang Ling, Ph.D.

Concurring Reviewers: Kun Jin, Ph.D., Team Leader
Jim Hung, Ph.D., Director

Medical Division: Division of Neuropharm (HFD -120)

Clinical Team: Christopher Breder, M.D., Ph.D.
Ronald Farkas, M.D., Ph.D., Team Leader

Project Manager: Fannie Choy

Table of Contents

LIST OF TABLES	3
LIST OF FIGURES	3
1 EXECUTIVE SUMMARY	4
2 INTRODUCTION	4
2.1 OVERVIEW	4
2.2 DATA SOURCES	5
3 STATISTICAL EVALUATION	5
3.1 DATA AND ANALYSIS QUALITY	5
3.2 EVALUATION OF EFFICACY	5
3.2.1 <i>Study Design and Endpoints</i>	5
3.2.2 <i>Statistical Methodologies</i>	7
3.2.3 <i>Patient Disposition, Demographic and Baseline Characteristics</i>	8
3.2.4 <i>Results and Conclusions</i>	9
3.3 EVALUATION OF SAFETY	15
4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS	16
4.1 GENDER, RACE, AGE, AND GEOGRAPHIC REGION	16
5 SUMMARY AND CONCLUSIONS	16
5.1 STATISTICAL ISSUES	16
5.2 COLLECTIVE EVIDENCE	16
5.3 CONCLUSIONS AND RECOMMENDATIONS	16

LIST OF TABLES

Table 1: Treatment Groups	6
Table 2: Demographic and Baseline Disease Characteristics.....	9
Table 3: Dystrophin-Positive Fibers Detected by IHC with MANDYS106.....	10
Table 4: Analysis Results of Change from Baseline in 6MWT.....	11
Table 5: Applicant's Result of 6MWT in Eteplirsen Compared to Historical Controls.....	14

LIST OF FIGURES

Figure 1. Overview of Study 201, Including Open-label Extension (Described as Study 202).....	6
Figure 2. LS Mean +/- SEM Change from Baseline in 6MWT - ITT Population.....	12
Figure 3: Historical Controls and Eteplirsen-Treated Cohort.....	13

1 EXECUTIVE SUMMARY

The data, overall, did not provide statistical evidence to support the efficacy of eteplirsen in patients who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping.

The only randomized controlled study submitted by the applicant, Study 201, can only be considered as exploratory because of study design and statistical analysis issues. In Study 201, patients were randomized to receive 50 or 30 mg/kg eteplirsen, or placebo. The study endpoints were assessed through Week 24. The statistical analysis plan of Study 201 did not include a method for statistical adjustment for testing multiple doses and/or multiple endpoints. The primary endpoint in Study 201 was the percent of dystrophin positive fibers in muscle biopsy tissue. The interpretation of the immunohistochemistry raw data is discussed in the clinical review. There was no nominally significant difference between eteplirsen 50 mg/kg, eteplirsen 30mg/kg and placebo for the 6MWT, which was the key clinical endpoint in Study 201.

The comparison of eteplirsen with historical controls, as proposed by the applicant in the open-label extension of Study 201 (called Study 202 by the applicant), is statistically uninterpretable, as this open-label extension did not have a prespecified statistical analysis plan, and had an inadequate control for bias. Among the potential sources of bias in the open-label extension of Study 201 are possible differences in various factors between eteplirsen-treated patients and the selected historical control cohort unaddressed by the applicant's attempt to match patients, the potential selection bias due to the *post-hoc* identification of the control cohort by the applicant, and other known sources of bias with the use of a historical control.

2 INTRODUCTION

2.1 Overview

Study 201 is the only randomized, double-blind, placebo-controlled study in this application. It was conducted at a single site in US in 12 subjects with genotypically confirmed DMD. Efficacy was assessed through the first 24 weeks of this study, while safety was assessed through Week 28. Upon completion of Study 201, all 12 patients were enrolled into an open-label extension study (Study 202) to continue receiving once-weekly treatment with eteplirsen. Study 202 was still ongoing at the time of NDA submission and interim study results were submitted for a cumulative 168 weeks of treatment, from Week 1 in Study 201 through the interim data cut at Week 140 in Study 202.

A historical control cohort was identified from 2 DMD patient registries for comparison to eteplirsen-treated patients in Study 201/202.

2.2 Data Sources

Materials reviewed for this application include the clinical study reports, raw and derived datasets, SAS codes used to generate the derived datasets and tables, protocols, statistical analysis plans, and documents of regulatory communications, which are located in the following directories: [\\CDSESUB1\evsprod\NDA206488\0001\m5\53-clin-stud-rep\535-rep-effic-safety-stud\dmd-51](#) and [\\CDSESUB1\evsprod\NDA206488\0006\m5\datasets](#).

3 STATISTICAL EVALUATION

3.1 Data and Analysis Quality

The key clinical efficacy endpoint results were reproduced by this reviewer from the raw data. Documentation of statistical analysis methods was included with sufficient details for this reviewer to reproduce the applicant's key efficacy results.

3.2 Evaluation of Efficacy

3.2.1 Study Design and Endpoints

The first patient was enrolled in Study 201 on July 18, 2011 and the study was completed on February 29, 2012. Protocol 201 was amended 7 times, 3 of them were implemented after the study was initiated and the last version was dated January 07, 2012. In Amendment 6 (dated November 04, 2011), the protocol changed the endpoint of 6-Minute Walk Test (6MWT) from exploratory endpoint to a secondary endpoint. In Amendment 7 (dated January 07, 2012), the duration of the study was extended from 24 to 28 weeks. The efficacy analyses were only specified in the statistical analysis plan (SAP), dated February 20, 2012.

Study 201 was not designed as a clinical efficacy study and not powered for efficacy analysis. The primary endpoint was the percent of dystrophin positive fibers as measured in muscle biopsy tissue, i.e., a biomarker. The key clinical secondary endpoint, 6MWT, was specified midway through the trial and the analyses were not specified until the trial was close to completion.

Study Design

This is a randomized, single-center, double-blind, placebo-controlled, multiple-dose study to assess the efficacy, safety, tolerability, and PK of once-weekly i.v. infusions of eteplirsen in subjects with genotypically confirmed DMD with an appropriate genetic lesion. Eligible subjects were randomized to receive 50 or 30 mg/kg eteplirsen or placebo, then placebo subjects were further randomized to 1 of 2 groups to create 4 treatment groups as shown in Table 1. Groups 1 and 2 received 50 or 30 mg/kg eteplirsen once a week for 28 weeks. Group 3a received placebo once a week for 24 weeks followed by 50 mg/kg eteplirsen for 4 weeks, and Group 3b received

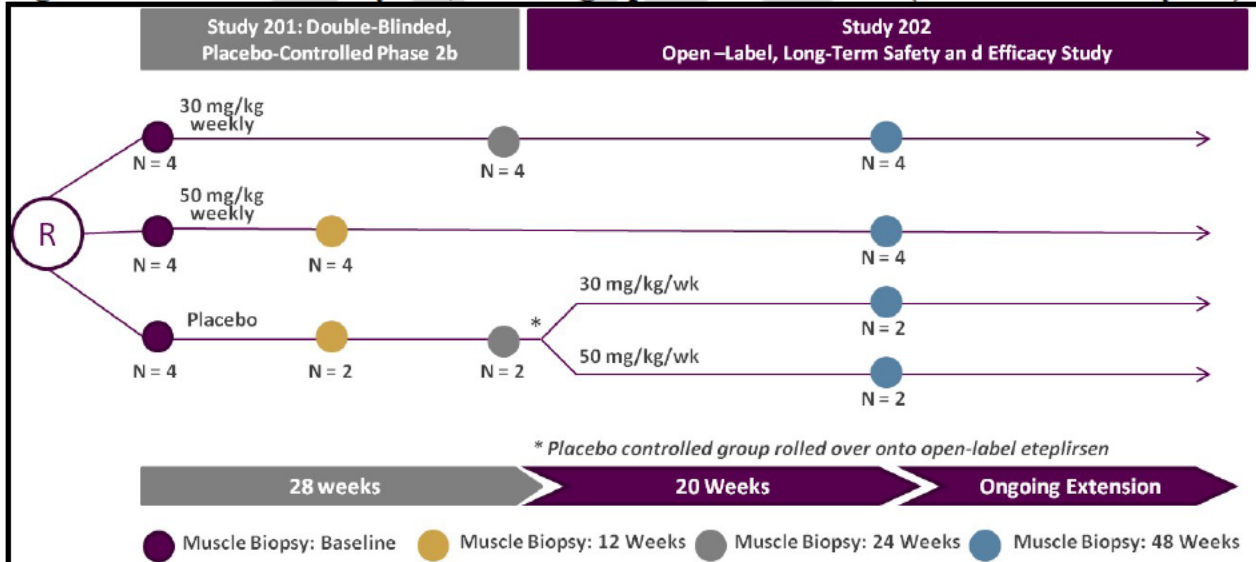
placebo once a week for 24 weeks followed by 30 mg/kg eteplirsen for 4 weeks. Beginning Week 25, all parties were aware that all subjects were receiving either 50 or 30 mg/kg eteplirsen.

Table 1: Treatment Groups

Group	N	Treatment
1	4	50 mg/kg eteplirsen IV once weekly for 28 weeks
2	4	30 mg/kg eteplirsen IV once weekly for 28 weeks
3a	2	Placebo IV for 24 weeks then 50 mg/kg eteplirsen for 4 weeks
3b	2	Placebo IV for 24 weeks then 30 mg/kg eteplirsen for 4 weeks

All patients underwent muscle biopsies at baseline for analysis of exon skipping and dystrophin expression. Repeat biopsies were performed at Week 12 for patients in Group 1 and Group 3a and at Week 24 for patients in Group 2 and Group 3b. Efficacy was assessed through the first 24 placebo-controlled weeks of this study, while safety was assessed through Week 28. Upon completion of this study, all 12 patients were rolled into an open-label extension (called Study 202 by the applicant) to continue receiving once-weekly treatment with eteplirsen for additional 212 weeks. In the open-label extension, all patients underwent a third muscle biopsy from the deltoid muscle at Week 20 and optionally a fourth muscle biopsy at approximately Week 140.

Figure 1. Overview of Study 201, Including Open-label Extension (Described as Study 202)



Efficacy Endpoints

Primary Efficacy Endpoint:

The primary efficacy endpoint is the change from baseline in percent of dystrophin positive fibers as measured in the muscle biopsy tissue using immunohistochemistry (IHC) at Week 12 for groups 1 and 3a and at Week 24 for groups 2 and 3b.

Key Efficacy Endpoints:

1. Changes from Baseline in CD3, CD4, and CD8 lymphocyte counts in muscle biopsy tissue at Week 12 for groups 1 and 3a and at Week 24 for groups 2 and 3b.
2. Changes from Baseline to Week 24 in 6-Minute Walk Test (6MWT).

The following clinical assessments were described as exploratory endpoints in the protocol (Amendment 7, dated 07 January 2012), but are included as key secondary endpoints together with 6MWT in the SAP (dated February 20, 2012):

- Timed 4 Step Test.
- Maximum voluntary isometric contraction test (MVICT) to measure elbow flexion and extension, knee flexion and extension, and grip strength.
- Timed 10-meter run from the North Star Ambulatory Assessment (NSAA).
- NSAA total score.

There is no clear description of hierarchal ordering among all those secondary endpoints. In the open-label extension (described as Study 202) only 6MWT is included as primary clinical endpoint.

3.2.2 Statistical Methodologies

Testing and summary statistics of all efficacy endpoints will combine placebo subjects into a single group. Some efficacy assessments including 6MWT were performed on Days 1 and 2 of the Week 1 (baseline), Week 12, and Week 24 visits and once at the Week 4, 8, 16, and 20 visits. On those visits where 2 tests were performed, the maximum/best observed value is used for the primary analysis. If data for any one visit day are missing, then the non-missing value from the same visit is used.

Efficacy Analysis Population

The efficacy analysis set is the Full Analysis Set (FAS), consisting of all subjects randomized into the study who received any amount of study drug.

Statistical Analysis Method

For this exploratory study, all statistical analyses are conducted at two-sided alpha level of 0.05. No multiplicity adjustment was specified for testing multiple doses and/or multiple endpoints, so all p-values are exploratory only.

The primary efficacy endpoint, the change from baseline in percent of dystrophin positive fibers, was analyzed by comparing the 50 mg/kg eteplirsen treatment group at Week 12 to the combined placebo treatment group, and the 30 mg/kg eteplirsen treatment group at Week 24 to the combined placebo treatment group, using the ANCOVA for ranked data with Baseline values and duration of DMD as covariates.

The analysis of changes from baseline to Week 24 in the clinical assessment parameters (6MWT, Timed 4 Step Test, MVICT, Timed 10-meter run, and NSAA total score) was based on a restricted maximum likelihood (REML)-based mixed model repeated measures (MMRM) with treatment (placebo, 30 mg/kg, 50 mg/kg), time, and treatment-by-time interaction terms as fixed effects, subject nested within treatment as random effects, with the Baseline value and time since DMD diagnosis as covariates. A first-order autoregressive (AR1) covariance structured matrix is used. The treatment comparison is made between each of the active treatments and placebo. The same MMRM analysis described above would be repeated to compare the combined eteplirsen groups to placebo.

If there was strong evidence suggesting that data for any endpoint deviated from normal distribution, then ANCOVA for ranked data was to be utilized.

3.2.3 Patient Disposition, Demographic and Baseline Characteristics

Patients were recruited for this study nationwide across the US. A total of 12 patients were randomized and all patients received scheduled infusions of study medication and completed the study as planned. All patients were 7 to 10-year old male and, except for one patient of Asian descent, all were white. The time since DMD diagnosis ranged from 18 to 112 months, with a median duration of 57 months. Numerically, there appears to be some imbalance in baseline 6MWT among the treatment groups (Table 2).

Table 2: Demographic and Baseline Disease Characteristics

Parameter		Placebo to Eteplirsena N = 4	Eteplirsena			All Patients N = 12
			30 mg/kg N = 4	50 mg/kg N = 4	All Eteplirsena N = 8	
Age	Mean	8.5	9.3	8.5	8.9	8.8
	Median	8.5	9	8.5	9	9
	Min, Max	7, 10	9, 10	7, 10	7, 10	7, 10
Mutation, n (%)	45-50	0	2 (50.0)	1 (25.0)	3 (37.5)	3 (25.0)
	48-50	0	1 (25.0)	0	1 (12.5)	1 (8.3)
	49-50	3 (75.0)	0	2 (50.0)	2 (25.0)	5 (41.7)
	50	1 (25.0)	0	0	0	1 (8.3)
	52	0	1 (25.0)	1 (25.0)	2 (25.0)	2 (16.7)
6MWT, meters	Mean	394.5	355.3	396	375.6	381.9
	Median	379	359	395	380.5	380
	SD	42.25	74.78	26.61	56.34	50.92
	Min, Max	364, 456	261, 442	365, 429	261, 442	261, 456
Time since DMD diagnosis, months	Mean	50.3	52.5	66.5	59.5	56.4
	Median	51	57	68	57	57
	SD	13.74	14.06	44.29	31.33	26.4
	Min, Max	36, 63	32, 64	18, 112	18, 112	18, 112

^a Includes both 30 mg/kg and 50 mg/kg

Source: Table 10-2 and 10-3 of the CSR.

3.2.4 Results and Conclusions

3.2.4.1 Analyses of the Primary Endpoint

The following analyses were based the fiber data derived by the applicant. The validity of the immunohistochemistry (IHC) raw data is beyond the scope of this review, and is addressed in the clinical review, to which the reader is referred for interpretation of the IHC results.

There was no statistically significant difference between the 50 mg/kg eteplirsena group and placebo at Week 12 ($p=0.958$; Table 3). At Week 24, the mean percentage of dystrophin-positive muscle fibers was higher in the eteplirsena 30 mg/kg group than the placebo. Patients treated with 30 mg/kg eteplirsena demonstrated 23% increase in the mean percentage of dystrophin positive fibers from baseline to Week 24. There appeared to be no increases from baseline in placebo patients. The nominal p value (0.002) for the comparison between eteplirsena 30 mg/kg group and the placebo group can only be considered exploratory, as there was no plan to control the type 1 error due to multiple comparisons, and the other comparison between 50mg/kg and placebo in Study 201 was negative.

Table 3: Dystrophin-Positive Fibers Detected by IHC with MANDYS106

Time point		Placebo	30 mg/kg Eteplirsen N = 4	50 mg/kg Eteplirsen N = 4
Baseline	Mean	15.64	18.19	11.00
	Median	15.58	17.80	11.51
	SD (SE)	10.742 (5.371)	5.501 (2.751)	4.668 (2.334)
	Min, Max	3.2, 28.2	11.9, 25.3	5.4, 15.6
On-Treatment	Mean	11.59	41.14	11.79
	Median	9.44	38.77	11.81
	SD (SE)	7.130 (3.565)	10.097 (5.049)	4.456 (2.228)
	Min, Max	5.7, 21.7	32.7, 54.3	6.4, 17.2
Change from Baseline	Mean	-4.05	22.95	0.79
	Median	-6.13	23.46	2.52
	SD (SE)	5.834 (2.917)	5.792 (2.896)	7.099 (3.549)
	Min, Max	-8.5, 4.5	15.9, 29.0	-9.3, 7.4
	p-value ^a		0.002	0.958

Source: CSR Table 11-1 and Table 14.2.1.1.2, confirmed by FDA reviewer.

^aBased on ANCOVA model for ranked data with treatment (placebo, 30 mg/kg, 50 mg/kg) as a fixed effect and baseline value and time since DMD diagnosis as covariates.

3.2.4.2 Analyses of 6MWT

As shown in Table 4, placebo-treated patients experienced a mean decline of 17.3 meters in 6MWT from baseline to Week 24, while patients in the 30 and 50 mg/kg eteplirsen groups showed mean declines of 134.8 and 2.3 meters, respectively. ANCOVA for ranked data showed no nominally significant differences between the treatment groups. The result of the MMRM analysis showed a nominally statistically significant difference between the placebo and 30 mg/kg eteplirsen groups, in favor of placebo (p=0.026; Table 4).

Table 4: Analysis Results of Change from Baseline in 6MWT

	Placebo	30mg/kg Eteplirsen N = 4	30mg/kg Eteplirsen mITT N = 2	50mg/kg Eteplirsen N = 4
Baseline				
Mean	394.5	355.3	407	396
Median	379	359	407	395
SD(SE)	42.25(21.12)	74.78(37.39)	49.50(35.00)	26.61(13.30)
Min, Max	364, 456	261, 442	372, 442	365, 429
Week 24				
Mean	377.3	220.5	394.5	393.8
Median	377.5	204	394.5	403.5
SD (SE)	19.00 (9.50)	203.14 (101.57)	51.62 (36.50)	53.67 (26.84)
Min, Max	354, 400	43, 431	358, 431	325, 443
Change at Week 24				
Mean	-17.3	-134.8	-12.5	-2.3
Median	-12	-116	-12.5	1.5
SD (SE)	28.06 (14.03)	144.71 (72.36)	2.12 (1.50)	29.89 (14.95)
Min, Max	-56, 11	-296, -11	-14, -11	-40, 28
treatment effect*		-102.4		25.6
95% CI *		(-192.2, -12.5)		(-62.7, 113.8)
P-value *		0.026		0.563

*Based on mixed model repeated measures (MMRM).

Source: Table 14.2.5.2.1 and Table 14.2.5.2.2 of Study 201 CSR,

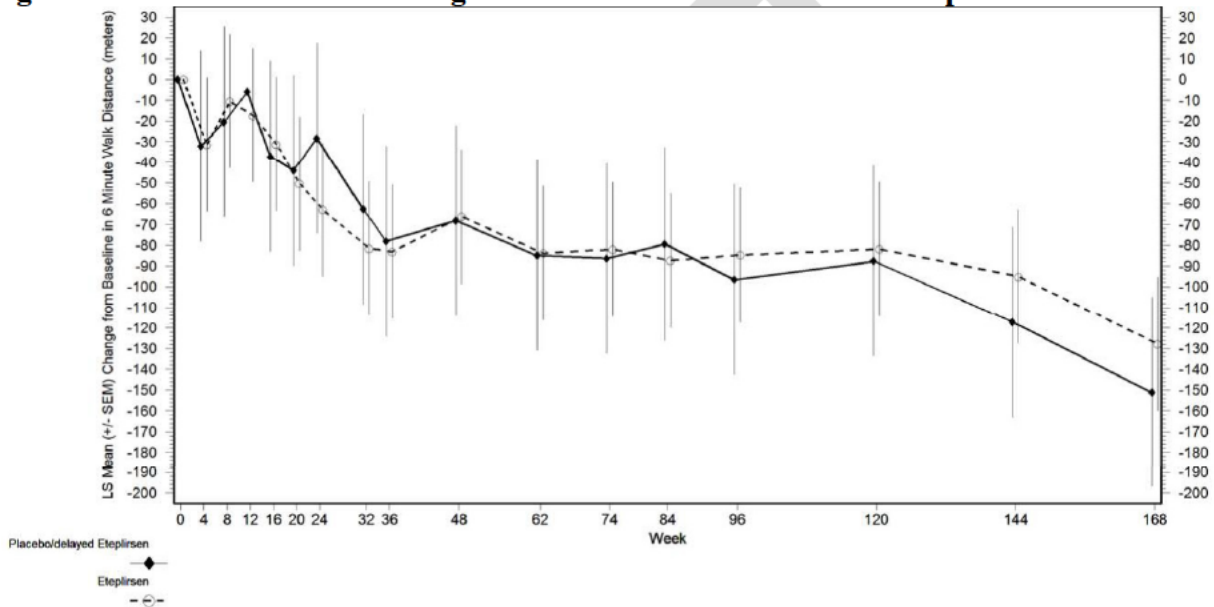
The applicant stated that the large decline in the 30 mg/kg eteplirsen group was attributable to Patients 009 and 010, who showed signs of rapid disease progression within weeks after enrollment. Therefore, the applicant conducted *post-hoc* analyses using Modified Intent-to-Treat (mITT) Population which excluded those 2 patients. For the mITT population, the mean change from baseline to Week 24 in MWT was a decline of 12.5 meters for the 30 mg/kg eteplirsen group. Both ANCOVA on ranked data and the MMRM analysis showed no nominally significant differences between the treatment groups in mITT.

The mITT population was not pre-specified in the SAP. Moreover, the mITT was defined based on the outcome data (instead of enrollment criteria or baseline character). Therefore, analysis on the mITT population could be misleading.

3.2.4.3 Analyses of the open-label extension study (described by the applicant as Study 202)

The 6MWT at Week 168 was compared between the combined eteplirsen group and placebo/delayed eteplirsen group. Analyses on ITT population did not achieve nominal statistical significance ($p=0.68$ by MMRM). The changes from baseline in 6MWT by assessment week for the combined eteplirsen group and placebo/delayed eteplirsen group are shown in Figure 2.

Figure 2. LS Mean \pm SEM Change from Baseline in 6MWT - ITT Population



Source: Figure 14.2.5.2.2.1 of Study 202 CSR.

3.2.4.4 Comparison against Historical Controls

Historical Control Cohort

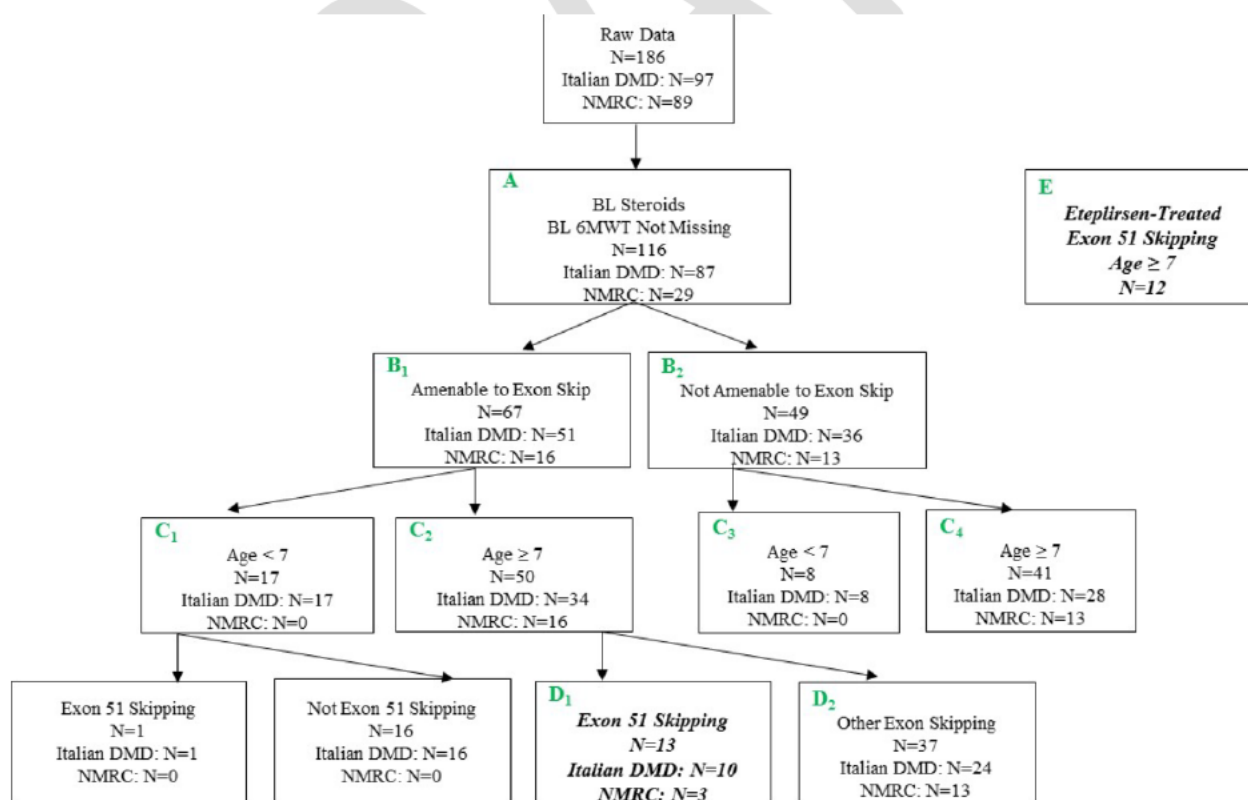
The comparison of eteplirsen with historical controls was not part of an adequate and well-controlled study. The applicant obtained historical data after observations were made for the eteplirsen patients. Historical data were obtained from 2 DMD patient registries (Italian DMD Registry and the Leuven Neuromuscular Reference Center – NMRC) for comparison to eteplirsen-treated patients. The following filters were applied to try to match patients in the historical control cohort:

1. Corticosteroid use at Baseline (use/non-use)
2. Sufficient longitudinal data for 6MWT available
3. Age ≥ 7 years
4. Genotype amenable to any exon skipping therapy
5. Genotype amenable to exon 51 skipping therapy

The Italian DMD registry is a longitudinal multicenter observational cohort study involving 11 tertiary neuromuscular centers in Italy. Patients were recruited between January 2008 and June 2010 and were to be followed for at least three years. The Italian DMD cohort contained the 6MWT results at Baseline (Month 0) and at Months 12, 24, and 36, with age and steroid use entered for each visit and with genotype information for 97 patients. Of these patients, 10 valid cases were identified based on applying the 5 filters.

The NMRC registry was an observational, single center, cohort study of DMD up to 17.5 years of age attending the NMRC between January 2007 and September 2012. The NMRC dataset contained 6MWT results at various time points, the patient's age and steroid use at the same time points, and genotype information for 89 patients. However, discrete visit designations (i.e., Baseline, Month 12, etc.) were not identified in the dataset. The first time points with non-zero meters on the 6MWT assessment for patients who were ≥ 7 years of age and on a steroid, were designated as the Baseline visit. Only 3 cases were identified based on applying filters (Figure 3).

Figure 3: Historical Controls and Eteplirsen-Treated Cohort



Source: Figure 1 of Study SR-15-031 CS.

Applicant’s Comparison of Eteplirsen with Historical Control

The results for 6MWT in eteplirsen-treated patients compared with historical controls matched on all 5 criteria mentioned above are shown in Table 5. The difference in LS mean change from baseline on 6MWT at 36 months was 141 meters. The nominal p-value reported by the applicant is not meaningful because the open label extension with historical control comparison was not an adequate and well-controlled study, for the reasons described below.

Table 5: Applicant’s Result of 6MWT in Eteplirsen Compared to Historical Controls

Patients Included	Groups Compared		Age	6MWT Baseline	6MWT Month 36**
HC + eteplirsen-treated, Steroid-Treated, Amenable to Exon 51 Skipping, ≥7 years old	HC	N	13	13	11
		Mean / LS Mean ^a (SE)	9.45 (0.403)	357.6 (18.51)	115.1 (33.54)
		Min, Max	7.3, 11.8	200, 458	
	eteplirsen-treated	N	12	12	12
		Mean / LS Mean ^a (SE)	9.41 (0.342)	363.2 (12.18)	256.4 (33.11)
		Min, Max	7.3, 11.0	256, 416	

* LS Mean for 6MWT Month 36 only

** LS Mean difference =141 and p=0.009 at month 36.

Source: Applicant’s analyses with output table modified by the reviewer.

Reviewer’s Discussion and Conclusion of the Historical Control Study

According to the ICH E10 guidance on Control Group and Related Issues in Clinical Trials, the major and well-recognized limitation of externally controlled (including historical control) trials is inability to control bias. The test group and control group can be dissimilar with respect to a wide range of observable and unobservable factors that could affect outcome. It may be possible to match the historical control group to the test group in observed factors but there is no assurance for any unobserved factors. “The lack of randomization and blinding, and the resultant problems with lack of assurance of comparability of test group and control group, make the possibility of substantial bias inherent in this design and impossible to quantitate.”

Because of the serious concern about the inability to control bias, the use of the external control design is restricted only to unusual circumstances.

1. ICH E10 states that “an externally controlled trial should generally be considered only when prior belief in the superiority of the test therapy to all available alternatives is so strong that alternative designs appear unacceptable...” However, such prior belief does not exist for eteplirsen.
2. ICH E10 states that “use of external controls should be limited to cases in which the endpoints are objective...” However, performance on the 6-minute walk test can be

influenced by motivation. Patients may not achieve maximal 6MWT due to concerns of falling or injury, or patients could try harder with encouragement and with the expectation that the drug might be effective.

3. Pocock's criteria¹ for acceptability of a historical control group require that "the methods of treatment evaluation must be the same," and "the previous study must have been performed in the same organization with largely the same clinical investigators." This is especially important when assessing endpoints such as 6MWT, in contrast to hard endpoints such as mortality. For this NDA, these requirements are not met.

Moreover, the historical control group was identified *post-hoc* in this NDA, leading to potential selection bias that cannot be quantitated. If a historical control is to be utilized, selection of the control group and matching on selection criteria should be prospectively planned without knowing the outcome of the drug group and control group.

Based on ICH E10, "a consequence of the recognized inability to control bias is that the potential persuasiveness of findings from externally controlled trials depends on obtaining much more extreme levels of statistical significance and much larger estimated differences between treatments than would be considered necessary in concurrently controlled trials." The success criteria for this historical control study were not discussed or pre-specified in the protocol.

Given all these concerns, including issues of comparability of eteplirsen-treated patients and historical control cohort patients, the fact that 6MWT is not a "hard" efficacy endpoint, the potential of selection bias due to the *post-hoc* identification of the control cohort by the applicant, and all the known pitfalls with the use of historical controls, the comparison of the eteplirsen with the historical control is not statistically interpretable.

3.3 Evaluation of Safety

Please see the clinical review.

¹ Pocock SJ. The combination of randomized and historical controls in clinical trials. *Journal of Chronic Diseases*. 1976; 29:175–188.

4 FINDINGS IN SPECIAL/SUBGROUP POPULATIONS

4.1 Gender, Race, Age, and Geographic Region

Subgroup analyses are not applicable as the study 201 was conducted at a single site in the US and all 12 patients were 7 to 10-year old male and, except for one patient of Asian descent, all were white.

5 SUMMARY AND CONCLUSIONS

5.1 Statistical Issues

Study 201 was designed as an exploratory study. No multiplicity adjustment was specified for testing multiple doses and/or multiple endpoints.

The sample sizes of both Study 201 and the historical control study are very small. The robustness of the study result is a concern since a single patient can change the results substantially. The interpretation of results is also difficult because the sample may not represent the DMD patient population at large. Small studies can be useful for hypothesis generating but usually do not have the ability to provide definitive evidence for a drug's effect.

5.2 Collective Evidence

In Study 201, there was no statistically significant difference between the 50 mg/kg eteplirsen group and placebo at Week 12 ($p=0.958$). Treatment with 30 mg/kg eteplirsen for 24 weeks increased the mean percentage of dystrophin-positive muscle fibers in DMD patients compared to placebo, however, the nominal p value (0.002) can only be considered exploratory due to the lack of multiplicity control.

The MMRM analysis of 6MWT at Week 24 in Study 201 showed a statistically significant difference between the placebo and 30 mg/kg eteplirsen groups, in favor of placebo ($p=0.026$). There was no statistically significant difference between the 50 mg/kg eteplirsen group and the placebo ($p=0.563$). These results must be considered as exploratory only.

The open-label extension with historical control is not statistically interpretable.

5.3 Conclusions and Recommendations

The data overall did not provide statistical evidence to support the efficacy in subjects who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping.

V. Summary of Clinical Pharmacology Findings

SUMMARY OF CLINICAL PHARMACOLOGY FINDINGS

NDA Number: 206488

Applicant Name: Sarepta Therapeutics, Inc.

Submission Dates: 08/20/2015, 11/02/2015

Brand Name: EXONDYS 51

Generic Name: Eteplirsen

Dosage Form: Injection

Dosage Strengths: Single use 2 mL vials containing 100 mg (50 mg/mL) of eteplirsen and single use 10 mL vials containing 500 mg (50 mg/mL) of eteplirsen

Proposed Indication: For the treatment of Duchenne muscular dystrophy (DMD) in patients who have a confirmed mutation of the *DMD* gene that is amenable to exon 51 skipping

OCP Division (s): DCP 1, DPM, Genomics and Targeted Therapy

Primary Reviewers: Atul Bhattaram, Ta-Chen Wu, Bart Rogers

Team Leaders: Kevin Krudys, Angela Men, Christian Grimstein

Natural History of DMD

Findings from the analyses of natural history data are included in the clinical division memo.

Pharmacokinetics:

- In general, dose-proportionality and linearity in PK properties may be concluded following weekly doses of 0.5~20 mg/kg in Phase 1 dose-ranging study and 30 and 50 mg/kg in efficacy trials. There was insignificant drug accumulation following weekly dosing across this dose range of 0.5~50 mg/kg.
- Following single or multiple IV infusion, the peak plasma concentrations (C_{max}) of eteplirsen occurred near the end of infusion and plasma concentration-time profiles of eteplirsen were generally similar and showed multi-phasic decline; the majority of drug elimination occurred within 24 hours.
- Plasma protein binding of eteplirsen in human is relatively low, ranging 6.1~16.5% and is independent of concentration studied.
- Distribution or cellular uptake of eteplirsen into peripheral tissues is supported by the volume of distribution (V_{ss}) values obtained following single or multiple doses (e.g., approximately 601 mL/kg after 30 mg/kg/week doses in Study 201).
- Eteplirsen was found to be metabolically stable in vitro with no evidence of metabolism or metabolite formation.
- The 30 and 50 mg/kg/wk doses studied in the clinical trials resulted in 64.1% and 69.4% of mean percent of dose excreted in the urine, total clearance of eteplirsen of 339 and 319 mL/hr/kg, and renal clearance of 221 and 234 mL/hr/kg (in the range of 115~125 mL/min), respectively. Elimination t_{1/2} was approximately 3.2~3.8 hours on average for the weekly 30 and 50 mg/kg doses.
- The inter-subject variability of eteplirsen was in the range of 20~55% for exposure measures (C_{max} and AUCs) as well as other key PK parameters.

Intrinsic factors:

Mutations Amenable to Exon 51 Skipping:

The sponsor has studied six different DMD mutations amenable to exon-51 skipping therapy. Additional DMD mutations (e.g. 19-50, 52-63) are known to exist, however they are ultra-rare (1-2 subjects in the Leiden database). If ultimately found to be safe and effective, eteplirsen should be approved for all mutations amenable to skipping of exon-51. While there may be some differences in functionality of the exon-51 skipped transcripts; restoring the reading frame to produce dystrophin even if it may be different between DMD mutations is warranted.

Extrinsic factors:

Drug-Drug Interaction (DDI)

In vitro studies:

Eteplirsen is expected to have a low potential for DDI in humans based on results of in vitro investigation on microsomal metabolism, plasma protein binding, inhibition or induction of major CYP isozymes or major drug transporters at the concentration range studied for clinical dosing regimen:

- Eteplirsen had insignificant inhibitory effects for CYP2B6, CYP2C8, CYP2D6, CYP3A4/5, CYP1A2, CYP2C9, or CYP2C19 in human liver microsomes. There was no metabolism-dependent inhibition observed with any of the CYPs tested.
- Eteplirsen at the concentration range studied did not show significant enzyme inducing capability for CYP1A2, CYP2B6, and CYP3A4 in human hepatocytes.
- Eteplirsen is not a substrate and/or an inhibitor of major human drug transporters OAT1, OAT3, OCT1, OCT2, OATP1B1, OATP1B3, P-gp, BCRP, MRP2, and BSEP in transfected CHO cells, Caco-2 monolayers, or inside-out human membranes.