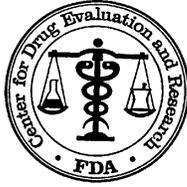


**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

206488Orig1s000

STATISTICAL REVIEW(S)



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

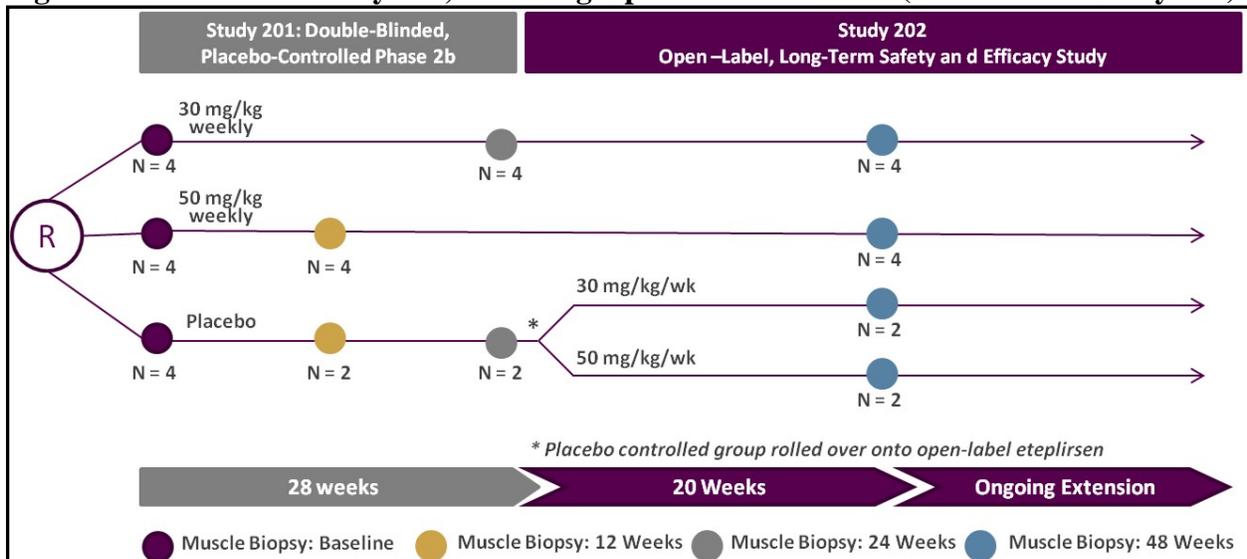
placebo once a week for 24 weeks followed by 30 mg/kg eteplirsen for 4 weeks. Beginning Week 25, all parties were aware that all subjects were receiving either 50 or 30 mg/kg eteplirsen.

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			All Patients
		Eteplirsena ^a	30 mg/kg	50 mg/kg	All Eteplirsena	
		N = 4	N = 4	N = 4	N = 8	N = 12
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,

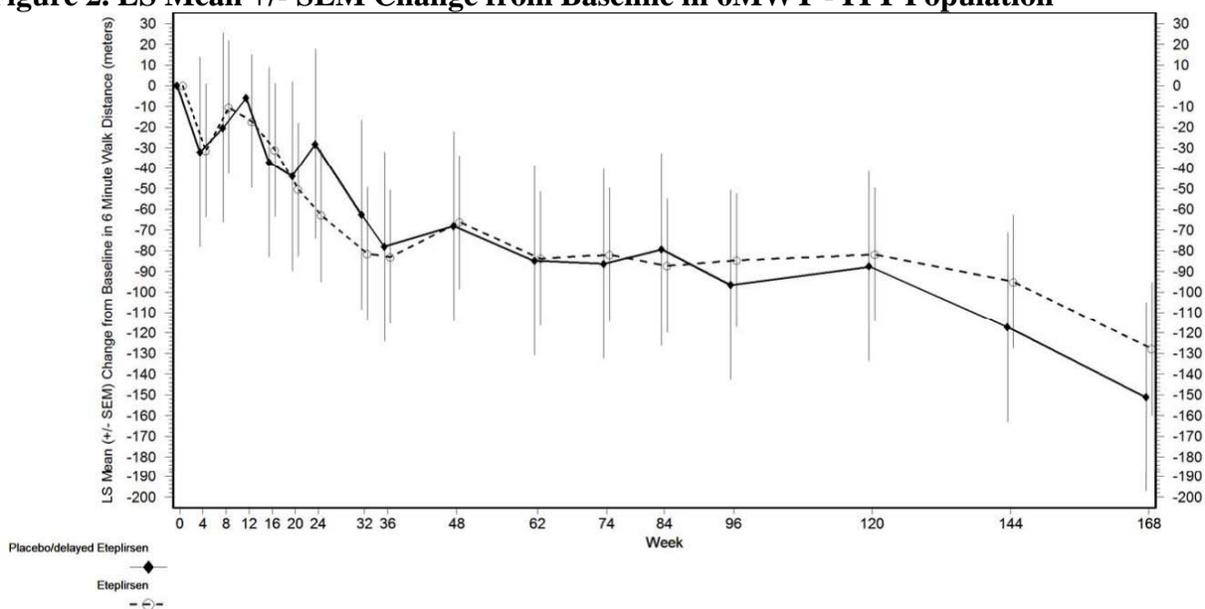
The applicant stated that the large decline in the 30 mg/kg eteplirsen group was attributable to Patients 009 and 010, who showed signs of rapid disease progression within weeks after enrollment. Therefore, the applicant conducted *post-hoc* analyses using Modified Intent-to-Treat (mITT) Population which 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

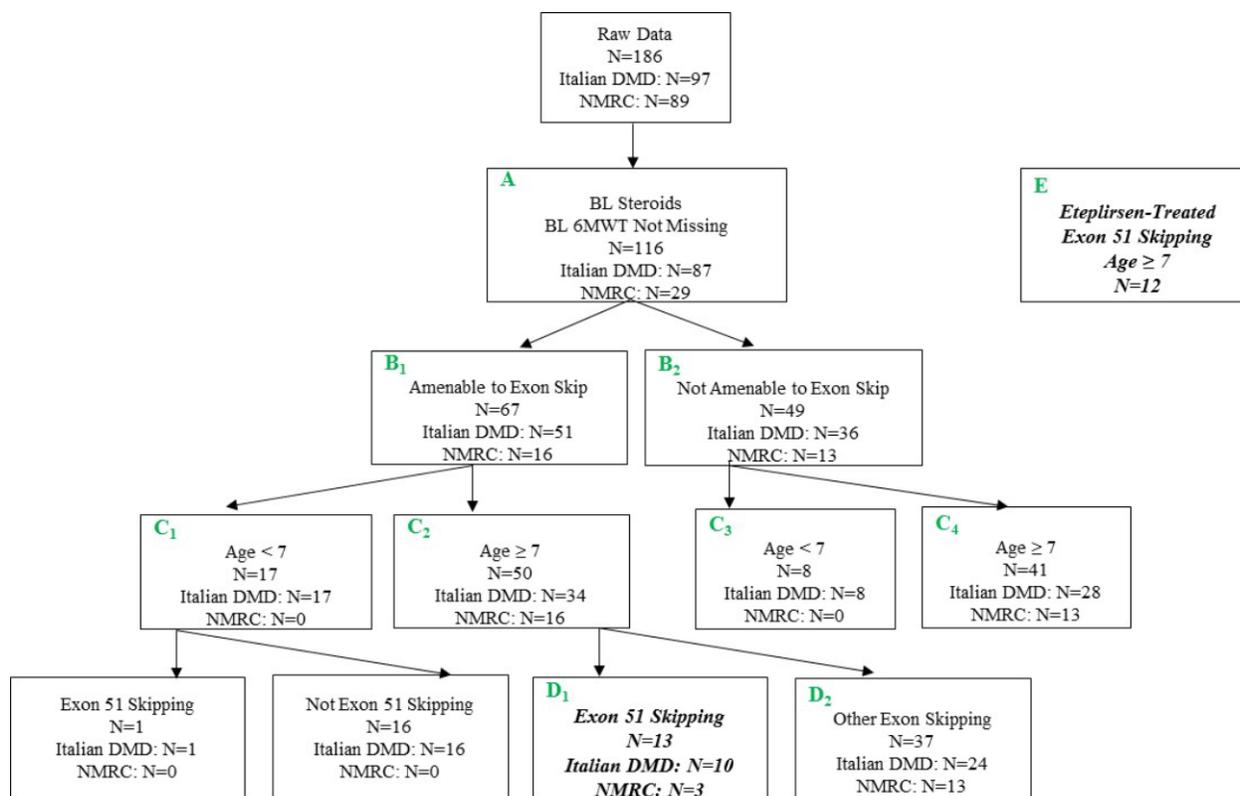
The comparison of eteplirsen with historical controls was not part of an adequate and well-controlled study. The applicant obtained historical data after observations were made for the eteplirsen patients. Historical data were obtained from 2 DMD patient registries (Italian DMD Registry and the Leuven 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.

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

XIANG LING
05/03/2016

KUN JIN
05/03/2016
I concur with the review.

HSIEN MING J HUNG
05/03/2016



US 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

Biometrics Division: VI

NDA No.:	206488
DATE RECEIVED BY OB:	September 28, 2015
DRUG NAME:	Eteplirsen,
SPONSOR:	Sarepta
INDICATION:	For treatment of some mutations that cause Duchenne muscular dystrophy (DMD), a genetic degenerative muscle disease
DOSAGE FORM:	Injection
FILL CONFIGURATION:	2.0 mL; 10.0 mL
REVIEW FINISHED DATE:	January 4, 2016
CMC STATISTICAL REVIEWER:	Zhuang Miao, Ph.D
CMC REVIEWER:	Mariappan Chelliah, Ph.D

Secondary Reviewer:

Xiaoyu Dong, Ph.D., Mathematical Statistician, CDER/OTS/OB/DB VI

Concur:

Yi Tsong, Ph.D., Division Director, DBVI, CDER/OTS/OB/DB VI

Meiyu Shen, Ph.D., Team Leader, Mathematical Statistician, CDER/OTS/OB/DB VI

Distribution:

CDER/OTS/OB/DB VI/ Yi Tsong

CDER/OTS/OB/DB VI/ Meiyu Shen

CDER/OTS/OB/ Lillian Patrician

CDER/OPQ/ONDP/DNDPI/NDPBI Wendy Wilson

CDER/OPQ/ONDP/DNDPI/NDPBI Mariappan Chelliah

CDER/OPQ/OPRO/DRBPMI/RBPMBI Dahlia Woody

Table of Contents

<i>I. Executive summary</i>	5
<i>II. Purpose of the review</i>	6
<i>III. Sponsor’s Stability Analysis and results</i>	6
<i>IV. FDA Statistical reviewer’s Analyses</i>	6
IV.1 Stability analysis for Assay of 2.0 mL	7
IV.2 Stability analysis for Assay of 10.0 mL	8
IV.3 Stability Analysis for Impurity [REDACTED] (b) (4) of 2.0 mL	10
IV.4 Stability Analysis for Impurity [REDACTED] of 10.0 mL	11
<i>V. Conclusions and Recommendations</i>	13

List of Tables:

Table 1: FDA Statistics Reviewer’s Estimated Shelf Life for Assay of each Fill Configuration using Long-term Stability Data..... 5

Table 2: FDA Statistics Reviewer’s Estimated Shelf Life for Impurity (b) (4) of each Fill Configuration using Long-term Stability Data 5

Table 3: Summary of Sponsor’s Shelf Life Determinations..... 6

Table 4: Poolability Testing Results for Stability Data of Assay 2.0 mL under the Long-term Storage Conditions..... 7

Table 5: By-batch Stability Regression Model estimation under the Long Term Stability Data of Assay 2.0 ml..... 7

Table 6: Poolability Testing Results for Stability Data of Assay under the Long-term Storage Conditions for Fill Configuration 10.0 mL..... 8

Table 7: By-batch Stability Regression Model estimation under the Long Term Stability Data of Assay 10.0 ml..... 9

Table 8: Poolability Testing Results for Stability Data of Impurity (b) (4) under the Long-term Storage Conditions for Fill Configuration 2.0 mL 10

Table 9: By-batch Stability Regression Model estimation under the Long Term Stability Data of Assay 2.0 ml..... 10

Table 10: Poolability Testing Results for Stability Data of Impurity (b) (4) 10.0 mL under the Long-term Storage 12

Table 11: By-batch Stability Regression Model estimation under the Long Term Stability Data of Impurity (b) (4) 10.0 ml 12

Table 12: FDA Statistics Reviewer’s Estimated Shelf Life for Assay of each Fill Configuration using Long-term Stability Data..... 13

Table 13: FDA Statistics Reviewer’s Estimated Shelf Life for Impurity (b) (4) of each Fill Configuration using Long-term Stability Data 13

List of Figures:

Figure 1: Stability Plot for Assay of 2.0 mL under the Long-term Conditions 8
Figure 2: Stability Plot for Assay of 10.0 mL under the Long-term Conditions 9
Figure 3: Stability Plot for Impurity [REDACTED]^{(b) (4)} of 2.0 mL under the Long-term Conditions 11
Figure 4: Stability Plot for Impurity [REDACTED] of 10.0 mL under the Long-term Conditions 12

I. EXECUTIVE SUMMARY

The purpose of this review is to determine if the proposed (b) (4) months shelf life for the new drug is supported. The sponsor's proposed shelf life of (b) (4) months is not supported by the current stability data. At the long term condition of 5 ± 3 °C, "if the long-term and accelerated data show little change over time and little variability", the proposed retest period or shelf life can be up to one-and-a-half times as long as, but should not be more than 6 months beyond the shortest last observation time based on ICH Q1E guidance. The estimations of shelf life for the drug product of 2.0 ml and drug product 10.0 ml are 18 months and 18 months because the shortest observed time period is only up to 12 months for Assay for both fill configurations

We performed statistical analysis on long-term stability data of Assay and Impurity (b) (4) of NDA 206488 with two fill configurations. Our estimations on the shelf life for each fill configuration are summarized in Table 1 and Table 2 below.

Table 1: FDA Statistics Reviewer's Estimated Shelf Life for Assay of each Fill Configuration using Long-term Stability Data

Fill Configuration	Acceptance Criteria	Method	Estimated Shelf Life (months)
2.0 mL	90%-115%	Common-slope-different-intercept	18
10.0 mL	90%-115%	Common-slope-different-intercept	18

Table 2: FDA Statistics Reviewer's Estimated Shelf Life for Impurity (b) (4) of each Fill Configuration using Long-term Stability Data

Fill Configuration	Acceptance Criteria	Method	Estimated Shelf Life (months)
2.0 mL	(b) (4) %	Common-slope-different-intercept	18
10.0 mL	(b) (4) %	Different-slope-different-intercept	18

Based on FDA statistics reviewer's independent analysis on the long-term stability data of assay and impurity (b) (4) our conclusions are summarized below:

- If there is no significant change at the accelerated storage condition, a shelf life of 18 months for the drug product is supported.
- Large batch-to-batch variability in the release data for Assay 2.0 mL.
- The variability for Assay 10.0 mL is larger than the variability of Assay 2.0 mL.
- We recommend that the sponsor provides more stability data.

Please note that, the shelf life estimation is performed under the assumption that the time trend beyond 18 months remains the same. The Sponsor’s analysis is summarized in Section III. The detailed analyses are provided in Section IV.

II. PURPOSE OF THE REVIEW

On September 28, 2015, Office of New Drug Products requested the CMC statistics team in Office of Biostatistics to evaluate the sponsor’s shelf life estimation for NDA 206488. The sponsor proposed a (b) (4) months shelf life based on their statistical analysis. The ONDP reviewer requested the OB reviewer to conduct the analysis in order to determine if the proposed (b) (4) months shelf life for the new drug is supported.

III. SPONSOR’S STABILITY ANALYSIS AND RESULTS

The sponsor did not provide the details of their stability analysis. In their report, they evaluated the attributes individually for all batches “because the determination from poolability testing was that it is not considered appropriate to combine the data from all batches”. For each attribute, shelf life was estimated by determining the earliest time at which the 95% confidence limit for the mean intersects the respective product specification acceptance criteria. The estimated shelf lives obtained from each test attribute for each lot are summarized in Table 3.

Table 3: Summary of Sponsor’s Shelf Life Determinations

Attribute	Shelf Life (months)					
	Batch Number					
	83GD-DR01	88GD-DR01	89GD-DR01	84GD-DQ01	85GD-DQ01	87GD-DQ01
Assay by IP-HPLC	64	40	102	36	40	39
Impurity (b) (4)	65	69	68	57	42	61

Reviewer’s Comment on the Sponsor’s Analysis:

1. For each fill configuration, the shortest last observation time is 12 months. It is not appropriate to estimate the shelf life more than one and a half time as long as 12 months or more than 6 months beyond 12 months for storage below room temperature if there is no significant change at accelerated condition based on ICH Q1E guidance.
2. The sponsor did not provide detailed information of the poolability test. The sponsor did not indicate if they used two-sided 95% confidence interval or one-sided 95% confidence interval.

IV. FDA STATISTICAL REVIEWER’S ANALYSES

There are two fill configurations, 2.0 mL and 10.0 mL. For each fill configuration, there are three primary stability batches. The data include 18 month data for one (1) batch and 12 month data for two (2) batches for each of the fill configurations (2-mL and 10-mL). Due to the deficiency of Sponsor’s analysis as pointed out before, we performed independent statistical analysis on the long-term stability data using ANCOVA method according to “Guidance for Industry Q1E

Evaluation of Stability Data”. The shelf life is estimated by the shortest time at which the two-sided 95% confidence limits of the mean value intercept with the acceptance criteria of 90% or 115%.

IV.1 Stability analysis for Assay of 2.0 mL

For Assay of 2.0 mL, we performed the poolability test based on the approach outlined in ICH Q1E guidance.

The P-values of batch and the interaction between time and batch are 0.4471×10^{-12} and 0.8472, respectively. The P-value for factor Batch is smaller than the significant level 0.25 and the P-value for Batch*Time is larger than the significant level 0.25. Thus, based on ICH Q1E guidance, the shelf life will be determined by a common-slope-different-intercept model.

Table 4: Poolability Testing Results for Stability Data of Assay 2.0 mL under the Long-term Storage Conditions

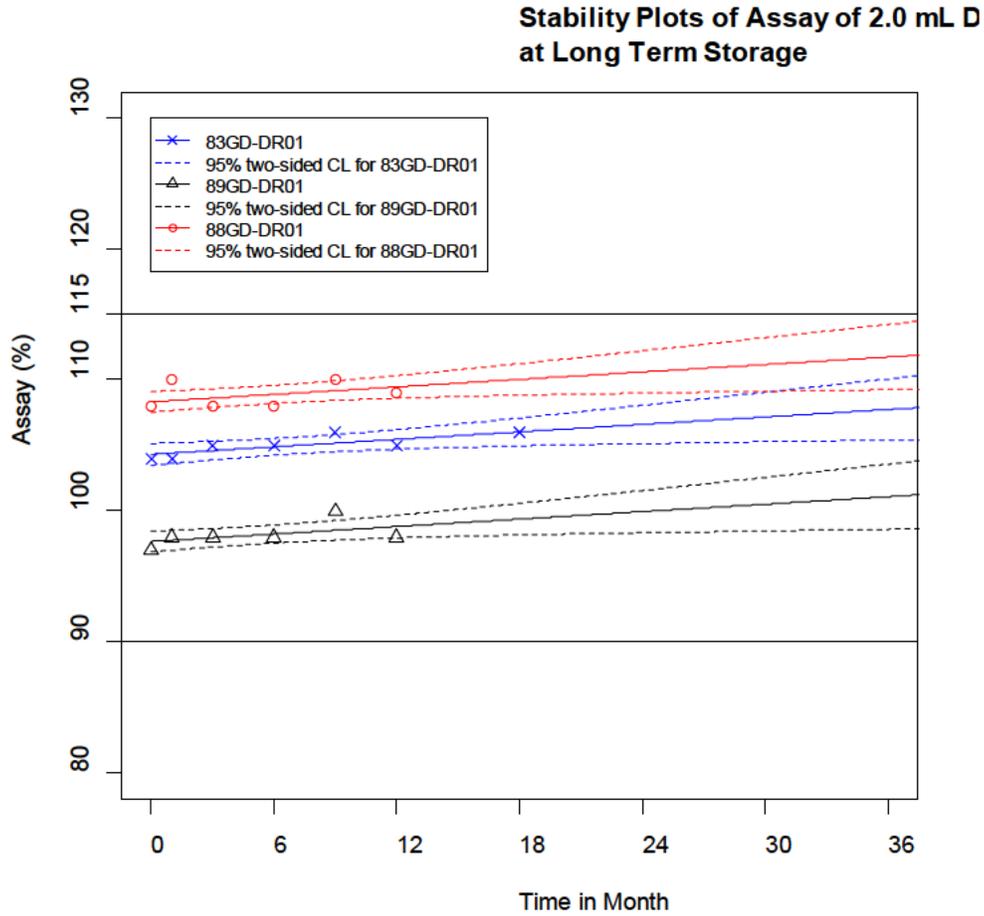
Variable	P-value	Significant Level
Batch	0.4471×10^{-12}	0.25
Batch*Time	0.8472	0.25

Table 5: By-batch Stability Regression Model estimation under the Long Term Stability Data of Assay 2.0 ml

Parameter	Estimated Intercept(Standard Error)	Estimated Slope(Standard Error)	Acceptance Criteria	Estimated Shelf Life (months)
83GD-DR01	104.335 (0.393)	0.095 (0.036)	90%-115%	>36 months
89GD-DR01	97.676 (0.445)			>36 months
88GD-DR01	108.342 (0.445)			>36 months

In Figure 2, the predicted mean values obtained by linear regression are shown in solid lines and the corresponding two-sided 95% confidence limits of the mean values are shown in dashed lines. The specified control limits are 90% and 115%. The shelf life is determined by the worst batch 88GD-DR01. The two-sided 95% confidence limit intercepts with the 115% acceptance limit at 40 months. The shelf life can only be extrapolated up to 6 months beyond the shortest last observed months for storage below room temperature if there are no significant changes under accelerated condition. Thus, the estimated shelf life for assay 2.0 mL is 18 months.

Figure 1: Stability Plot for Assay of 2.0 mL under the Long-term Conditions



IV.2 Stability analysis for Assay of 10.0 mL

For Assay of 10.0 mL, we performed the poolability test based on the approach outlined in ICH Q1E guidance.

The P-values of batch and the interaction between time and batch are 0.2434 and 0.6338, respectively. The P-value for factor Batch is smaller than the significant level 0.25 and the P-value for Batch*Time is larger than the significant level 0.25. Thus, based on ICH Q1E guidance, the shelf life will be determined by a common-slope-different-intercept model.

Table 6: Poolability Testing Results for Stability Data of Assay under the Long-term Storage Conditions for Fill Configuration 10.0 mL

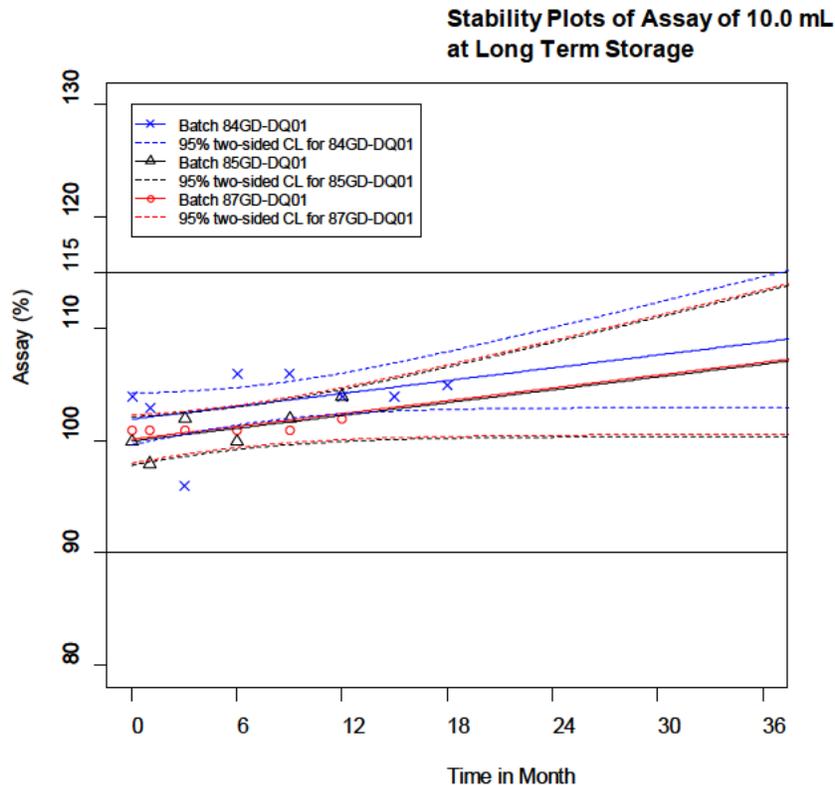
Variable	P-value	Significant Level
Batch	0.2434	0.25
Batch*Time	0.6338	0.25

Table 7: By-batch Stability Regression Model estimation under the Long Term Stability Data of Assay 10.0 ml

Parameter	Estimated Intercept (Standard Error)	Estimated Slope(Standard Error)	Acceptance Criteria	Estimated Shelf Life (months)
84GD-DQ01	101.977 (1.078)	0.190 (0.094)	90%-115%	36 months
85GD-DQ01	100.016 (1.204)			>36 months
87GD-DQ01	100.183 (1.204)			>36 months

In Figure 3, the predicted mean values obtained by linear regression are shown in solid lines and the corresponding two-sided 95% confidence limits of the mean values are shown in dashed lines. The specified control limits are 90% and 115%. The shelf life is determined by the worse batch 84GD-DQ01. The two-sided 95% confidence limit intercepts with the 115% acceptance limit at 36 months. The shelf life can only be extrapolated up to 6 months beyond the shortest last observed months for storage below room temperature if there are no significant changes under accelerated condition. Thus, the estimated shelf life for assay 10.0 mL is 18 months.

Figure 2: Stability Plot for Assay of 10.0 mL under the Long-term Conditions



IV.3 Stability Analysis for Impurity (b) (4) of 2.0 mL

For Impurity (b) (4) of 2.0 mL, we performed the poolability test based on the approach outlined in ICH Q1E guidance.

The p-values of batch and the interaction between time and batch are (b) (4) respectively. (b) (4) based on ICH Q1E guidance, the shelf life will be determined by a common-slope-different-intercept model.

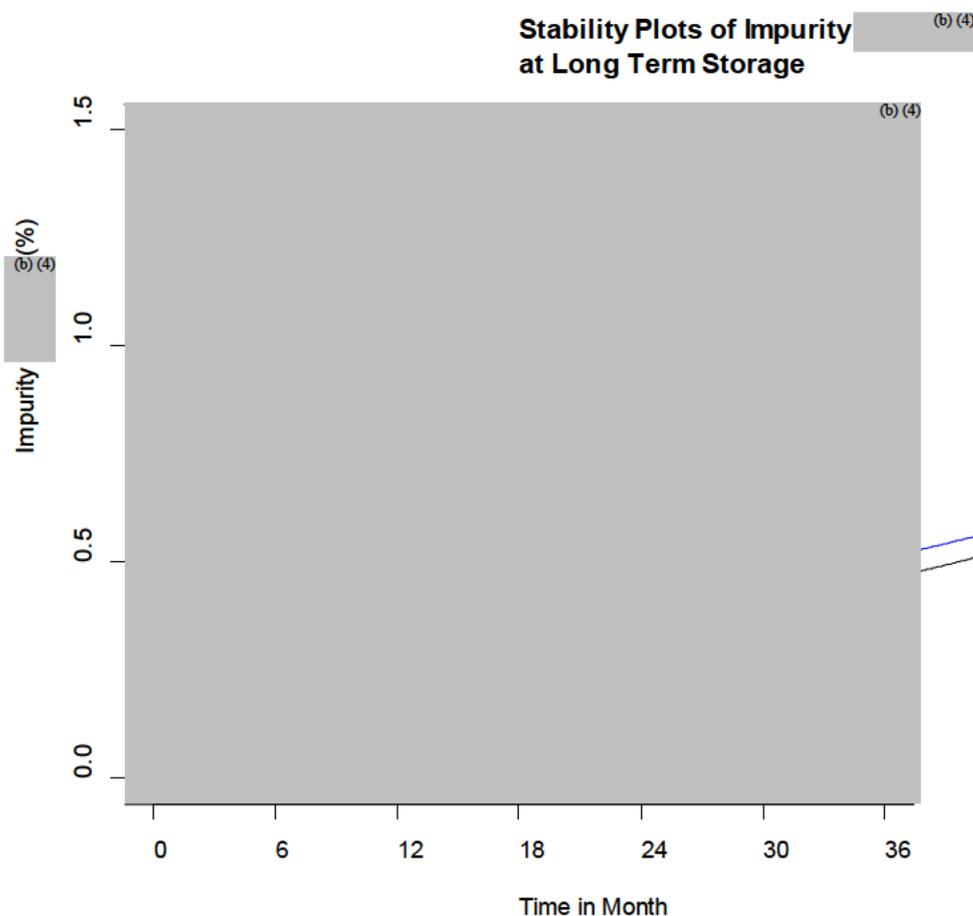
Table 8: Poolability Testing Results for Stability Data of Impurity (b) (4) under the Long-term Storage Conditions for Fill Configuration 2.0 mL

Variable	P-value	Significant Level
Batch	(b) (4)	(b) (4)
Batch*Time	(b) (4)	(b) (4)

Table 9: By-batch Stability Regression Model estimation under the Long Term Stability Data of Assay 2.0 ml

Parameter	Estimated Intercept(Standard Error)	Estimated Slope(Standard Error)	Acceptance Criteria	Estimated Shelf Life (months)
83GD-DR01	(b) (4)			>36 months
89GD-DR01	(b) (4)	(b) (4)	(b) (4) %	>36 months
88GD-DR01	(b) (4)			>36 months

Figure 3: Stability Plot for Impurity (b) (4) of 2.0 mL under the Long-term Conditions



In Figure 4, the predicted mean values obtained by linear regression are shown in solid lines and the corresponding one-sided 95% confidence limits of the mean values are shown in dashed lines. The specified control limit is (b) (4)%. The shelf life is determined by the worse batch. The one-sided 95% confidence limit intercepts with the (b) (4)% acceptance limit at >36 months. The shelf life can only be extrapolated up to 6 months beyond the shortest last observed months for storage below room temperature if there are no significant changes under accelerated condition. Thus, the estimated shelf life for impurity (b) (4) 2.0 mL is 18 months.

IV.4 Stability Analysis for Impurity (b) (4) of 10.0 mL

For Impurity (b) (4) of 10.0 mL, we performed the poolability test based on the approach outlined in ICH Q1E guidance.

The p-values of batch and the interaction between time and batch are (b) (4), respectively. (b) (4) based on ICH Q1E guidance, the shelf life will be determined by a different-slope-different-intercept model.

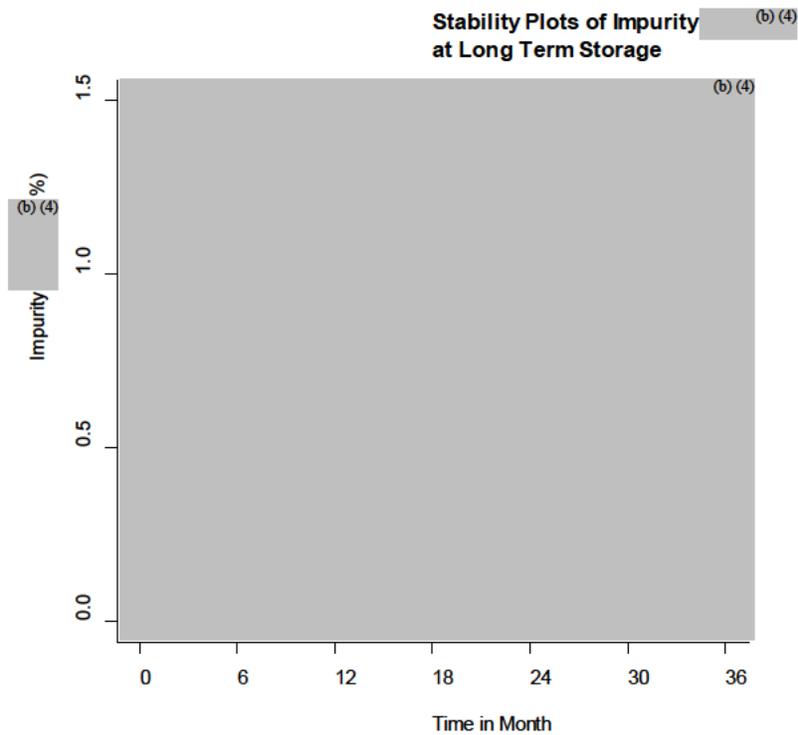
Table 10: Poolability Testing Results for Stability Data of Impurity (b) (4) 10.0 mL under the Long-term Storage

Variable	P-value	Significant Level
Batch	(b) (4)	(b) (4)
Batch*Time	(b) (4)	(b) (4)

Table 11: By-batch Stability Regression Model estimation under the Long Term Stability Data of Impurity (b) (4) 10.0 ml

Parameter	Estimated Intercept(Standard Error)	Estimated Slope(Standard Error)	Acceptance Criteria	Estimated Shelf Life (months)
84GD-DQ01	(b) (4)	(b) (4)	(b) (4) %	>36 months
85GD-DQ01	(b) (4)	(b) (4)		>36 months
87GD-DQ01	(b) (4)	(b) (4)		>36 months

Figure 4: Stability Plot for Impurity (b) (4) of 10.0 mL under the Long-term Conditions



In Figure 5, the predicted mean values obtained by linear regression are shown in solid lines and the corresponding one-sided 95% confidence limits of the mean values are shown in dashed lines. The specified control limit is (b)(4)%. The shelf life is determined by the worse batch 85GD-DQ01. The one-sided 95% confidence limit intercepts with the (b)(4)% acceptance limit at >36 months. The shelf life can only be extrapolated up to 6 months beyond the shortest last observed months for storage below room temperature if there are no significant changes under accelerated condition. Thus, the estimated shelf life for impurity (b)(4) 10.0 mL is 18 months.

V. CONCLUSIONS AND RECOMMENDATIONS

We performed stability analysis for assay value and impurity (b)(4) of fill configuration 2.0 mL and 10.0 mL. The sponsor's proposed shelf life of (b)(4) months is not supported by the current stability data. The estimations of shelf life for the drug product of 2.0 ml and drug product 10.0 ml are 18 months and 18 months. The analysis results are summarized in the following tables. In addition, we recommend that the sponsor submit more stability data.

Table 12: FDA Statistics Reviewer's Estimated Shelf Life for Assay of each Fill Configuration using Long-term Stability Data

Fill Configuration	Acceptance Criteria	Method	Estimated Shelf Life (months)
2.0 mL	90% -115%	Common-slope-different-intercept	18
10.0 mL		Common-slope-different-intercept	18

Table 13: FDA Statistics Reviewer's Estimated Shelf Life for Impurity (b)(4) of each Fill Configuration using Long-term Stability Data

Fill Configuration	Acceptance Criteria	Method	Estimated Shelf Life (months)
2.0 mL	(b)(4)%	Common-slope-different-intercept	18
10.0 mL		Different-slope-different-intercept	18

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

ZHUANG MIAO
01/05/2016

XIAOYU DONG
01/06/2016

MEIYU SHEN
01/06/2016

YI TSONG
01/07/2016

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

NDA Number: 206488

Applicant: Sarepta

Stamp Date: 6/26/15

Drug Name: eteplirsen

NDA/BLA Type: original NDA

On **initial** overview of the NDA/BLA application for RTF:

	Content Parameter	Yes	No	NA	Comments
1	Index is sufficient to locate necessary reports, tables, data, etc.	x			Links do not work
2	ISS, ISE, and complete study reports are available (including original protocols, subsequent amendments, etc.)	x			
3	Safety and efficacy were investigated for gender, racial, and geriatric subgroups investigated (if applicable).			x	
4	Data sets in EDR are accessible and do they conform to applicable guidances (e.g., existence of define.pdf file for data sets).	x			

IS THE STATISTICAL SECTION OF THE APPLICATION FILEABLE? ___yes___

If the NDA/BLA is not fileable from the statistical perspective, state the reasons and provide comments to be sent to the Applicant.

Please identify and list any potential review issues to be forwarded to the Applicant for the 74-day letter.

Content Parameter (possible review concerns for 74-day letter)	Yes	No	NA	Comment
Designs utilized are appropriate for the indications requested.	x			
Endpoints and methods of analysis are specified in the protocols/statistical analysis plans.	x			
Interim analyses (if present) were pre-specified in the protocol and appropriate adjustments in significance level made. DSMB meeting minutes and data are available.			x	
Appropriate references for novel statistical methodology (if present) are included.			x	
Safety data organized to permit analyses across clinical trials in the NDA/BLA.	x			
Investigation of effect of dropouts on statistical analyses as described by applicant appears adequate.			x	

File name: 5_Statistics Filing Checklist for a New NDA_BLA110207

STATISTICS FILING CHECKLIST FOR A NEW NDA/BLA

Xiang ling

8/2/2015

Reviewing Statistician

Date

Supervisor/Team Leader

Date

This is a representation of an electronic record that was signed electronically and this page is the manifestation of the electronic signature.

/s/

XIANG LING
08/05/2015

KUN JIN
08/19/2015