Agency Scientific Dispute – Appeal

Date: **July 18, 2016**

To: **G. Matthew Warren**
Director
Office of Scientific Integrity, FDA

From: **Ellis F. Unger, M.D. (initiator)**
Director
Office of Drug Evaluation-I
Office of New Drugs
Center for Drug Research and Evaluation
U.S. Food and Drug Administration

Re: **NDA #** 206488
Drug: eteplirsen (Exondys 51)
Applicant: Sarepta Therapeutics
Indication: Treatment of Duchenne muscular dystrophy (DMD) in patients who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping

1. Background

The Office of New Drugs within the Center for Drug Evaluation and Research (CDER) oversees regulation of new drugs, and is responsible for making regulatory decisions for approval/non-approval of new molecular entities. Within the Office of New Drugs, there are 6 sub-offices, including the Office of Drug Evaluation-I. The Office of Drug Evaluation-I oversees the Division of Neurology Products, which regulates drugs for the central and peripheral nervous systems, as well as drugs for muscular disorders. Typically, a new drug application (NDA) for a new molecular entity for a neurology indication is reviewed by the Division of Neurology Products in concert with review staff from other offices in CDER.[1] The regulatory decision is typically rendered by Office of Drug Evaluation-I, i.e., the signatory authority.

NDA 206488 for eteplirsen was reviewed by the Division of Neurology Products, and members of the review team reached the unanimous conclusion that the NDA should receive a *complete response* action. This view was shared by the Office of Biometrics, which performed the statistical review, as well as the Office of Clinical Pharmacology, which performed the pharmacology review. Dr. John Jenkins, Director, Office of New Drugs, also supports a *complete response* action for this NDA (verbal communication).

This memo is meant to explain the salient arguments around the scientific disagreement here; additional details are available in my memo recommending a complete response and Dr. Woodcock's memo recommending approval, and the reader is referred to those memoranda.

Disease Background:

[1] Reviews are typically provided by Office of New Drug Quality Assessment, Division of Medication Error Prevention and Analysis, Office of Biometrics, Office of Scientific Investigations, and others.

Duchenne muscular dystrophy is an X-linked recessive neuromuscular disorder caused by mutations of the dystrophin gene. These mutations disrupt the messenger ribonucleic acid (mRNA) reading frame, leading to the absence or near-absence of dystrophin protein in muscle cells. The disorder affects 1 in ~3,600 boys.

Dystrophin protein is thought to maintain the structural integrity of the muscle cell, cushioning it from the stress and strain of repeated contraction and relaxation. Absence of dystrophin leads to muscle damage, with replacement by fat and collagen. With progressive degeneration of skeletal muscle (including breathing muscles) and cardiac muscle, there is loss of physical function in childhood and adolescence, with premature death from respiratory and/or cardiac failure in the second to fourth decade.

No specific therapies are approved for DMD. Steroids are currently the cornerstone of management, widely believed to delay loss of ambulation and respiratory decline by several years.

Drug Background:

Eteplirsen is a phosphorodiamidate morpholino oligomer (PMO) designed to target the pre-mRNA transcripts of the dystrophin gene so that exon 51 is excluded, or skipped, from the mature, spliced mRNA. Theoretically, by restoring of the mRNA reading frame, a 'truncated' but nevertheless partially functional form of the dystrophin protein can be produced by muscle cells, delaying disease progression. Similar truncated dystrophin is found in a less severe form of muscular dystrophy, Becker Muscular Dystrophy (BMD). In essence, the drug is hoped to induce production of sufficient Becker-type dystrophin to slow the progression of the disease. This drug is specific for exon 51 mutations, a subset of the mutations that cause DMD. If approved, the drug would be indicated for ~13% of the overall DMD patient population. Eteplirsen has not received marketing authorization from any regulatory authority, and no similar drugs are approved.

Drug Development Background:

Three studies are germane to the issues here. Study 201 was a single-center, double-blind, placebo-controlled study in 12 patients with DMD. Patients were randomized (1:1:1) to eteplirsen 30 mg/kg/week, eteplirsen 50 mg/kg/week, or placebo (4 patients per group). After 24 weeks, the 4 patients originally randomized to placebo were re-randomized to eteplirsen 30 mg/kg/week (n=2) or eteplirsen 50 mg/kg/week (n=2). The trial was eventually extended to an open-label phase (Study 202) where all patients received eteplirsen, although investigators and patients remained blinded to dose. These patients have continued to receive eteplirsen for more than 4 years. This continuous study is referred to as Study 201/202. Study 301 is an externally controlled study where all patients are receiving open-label eteplirsen, 30 mg/kg, by weekly infusion. The study is ongoing and still accruing patients. Interim data were obtained from 13 patients in this study (see below).

The endpoints for these studies can be broadly divided into those that aim to show changes in physical performance, e.g., walking speed, rise time from the floor, muscle function; and those that aim to show effects on production of dystrophin in skeletal muscle – the surrogate endpoint. Dystrophin was quantified in this development program using two methods: Western blot and immunohistochemistry.

2. Description of How My Position Differs from the Center's Perspective

Dr. Janet Woodcock, Director, CDER, disagrees with some of the findings of the review team, and has reached the conclusion that the NDA should be approved. She finds that the data meet the standard for accelerated approval under 21 CFR 314. 510, based on the change in a surrogate endpoint of dystrophin protein production – a change she concludes is reasonably likely to predict clinical benefit. The disagreement is over the question of whether the findings on the dystrophin surrogate endpoint are reasonably likely to predict clinical benefit. The decision of *approval* vs. *complete response* hinges on this question.

a. Clinical/Statistical Efficacy

Accelerated Approval:

Dr. Woodcock has reached the conclusion that eteplirsen should receive accelerated approval based on a small effect on the surrogate endpoint of dystrophin production.

The relevant statutory and regulatory framework (section 506(c) of the FD&C Act and 21 CFR part 314, subpart H) states that a drug can receive accelerated approval if 3 factors are satisfied:

1. If the drug treats a serious or life-threatening disease or condition,
2. if FDA takes into account the severity, rarity, or prevalence of the condition and the availability or lack of alternative treatments, and
3. if the drug demonstrates an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit OR demonstrates an effect on an intermediate 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. As noted in section 506(c)(1)(B) of the FD&C Act, the evidence to support the concept "…that an endpoint is reasonably likely to predict clinical benefit may include epidemiological, pathophysiological, therapeutic, pharmacologic, or other evidence developed using biomarkers, for example, or other scientific methods or tools."

In terms of the prospect for accelerated approval for eteplirsen, DMD is clearly a serious, severe, and rare condition with no approved treatments; therefore, factors 1 and 2, above, are satisfied. There is no disagreement.

The critical issue is whether factor 3 is satisfied, and factor 3 can be subdivided into three parts: 1) whether the surrogate endpoint is appropriate for the disease; 2) whether there is substantial evidence of an effect on the surrogate endpoint; and 2) whether the effect demonstrated meets the test of being "reasonably likely" to predict clinical benefit. Importantly, there is no regulatory definition of "reasonably likely."

For the first part of factor 3, whether the surrogate endpoint is appropriate for the disease, the review team has agreed that the near-lack of dystrophin is the proximal cause of DMD, and that the level of dystrophin in skeletal muscle is an appropriate surrogate endpoint that could predict efficacy. There is no disagreement here.

The second part of factor 3 is whether an effect has been demonstrated; the legal standard is 'substantial evidence' based on adequate and well-controlled clinical investigations. Typically, such evidence would be two studies, both achieving a *p*-value < 0.05, but in some situations FDA has the flexibility to interpret data from a single trial, or a single trial with supporting evidence, as substantial evidence of effectiveness.[2] Dr. Woodcock believes that "…there is evidence from adequate and well-controlled trials, and supportive evidence, that exposure to eteplirsen increases dystrophin protein production in muscle cells." I agree that there is evidence from a single adequate and well controlled trial, Study 301, that eteplirsen induces dystrophin production in muscle cells, but do not agree that there is reliable quantitative evidence from the other trial, Study 201/202.

The third part of factor 3, the conclusion that the demonstrated effect is "reasonably likely" to predict clinical benefit, is where there is disagreement.

A. Are the Data on Dystrophin Protein Production from One or More Adequate and Well-Controlled Studies?

Dr. Woodcock cites 3 lines of evidence pertinent to the conclusion that eteplirsen increases dystrophin production:

1. Production of an appropriate mRNA transcript
2. Quantitative assessment of dystrophin content in muscle biopsies by Western blot
3. Semi-quantitative assessment of dystrophin in muscle tissue by immunohistochemistry

1. <u>Production of an appropriate mRNA transcript</u>

I agree that the applicant has shown expression of mRNA following treatment with eteplirsen. As noted by Dr. Woodcock, this finding establishes proof of concept, but does not by itself mean that there is increased dystrophin production.

2. <u>Quantitative assessment of dystrophin content in muscle biopsies by Western blot</u>

Western blot is a standard laboratory technique used to quantify proteins in body tissues. In Sarepta's development program, Western blot was used to assess dystrophin protein levels in skeletal muscle in Study 201, in Study 202 (again, these were Study 201 patients who were maintained on treatment), and finally in Study 301.

a. Study 201:

The original Western blot analyses from Study 201 were intended to show that dystrophin levels were greater in eteplirsen-treated patients than in patients in the placebo group, and analyses were planned to compare the effects of the lower vs. higher eteplirsen doses on dystrophin production. The Western blots submitted by the applicant for Study 201 were oversaturated, unreliable, and uninterpretable.

b. Study 202:

[2] See: "Guidance for Industry Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products;" May, 1998.

With FDA's assistance, the applicant improved the assays and performed repeat biopsies on 11 of 12 patients of the Study 201/202 patients at Week 180. These were to be compared to stored baseline (pre-treatment) samples; however, evaluable tissue was available for only 3 of the 11 patients. The baseline samples are germane to the determination of the treatment effect because the Week 180 biopsies showed only a small quantity of dystrophin (mean = 0.93% of normal). Thus, for the purpose of computing the *change* in dystrophin resulting from eteplirsen treatment, even small differences in the baseline level are critical.

As noted by Dr. Woodcock, the review team and I had concerns about these controls, leading us to conclude that Study 201/202 was not adequate and well controlled:

1. The goal was to assess the change in dystrophin with treatment, i.e., pre-treatment vs. post-treatment, but most of the baseline biopsies were obtained from subjects external to Study 201/202, who could differ in unknown ways from subjects in Study 201/202.

2. For all patients, the Week 180 biopsies were obtained from different muscles than the baseline biopsies, and studies of both normal human muscle and non-clinical DMD models have shown that dystrophin levels vary among muscles.

3. The baseline biopsies for the three subjects with Week 180 data had been stored for several years and the protein may have degraded, leading to a falsely low baseline value, and a greater apparent increase from baseline, accordingly.

Dr. Woodcock believes that "…these issues increase the uncertainty around the results, but do not necessarily render them an inadequate basis on which to draw a conclusion." She notes that the external control patients were similar in age and mutation site to the patients in Study 201/202. She found little difference between dystrophin results across different muscle groups, and little difference based on storage time, leading her to believe that these factors "…did not result in large differences in the findings."

Although I agree that these factors are not likely to lead to large differences, even small differences would affect the calculation of the *change* in dystrophin at Week 180, because the Week 180 values were quite small (mean only 0.93% of normal). At issue is how much of the dystrophin detected at Week 180 was newly produced, vs present at baseline. For example, a difference in the baseline level of only 0.30%, although minute, is substantial compared to 0.93%.

Dr. Woodcock notes that at Week 180, 2 subjects had dystrophin levels between 2 and 3%, 2 had a level between 1 and 2%, and 2 had a level of ~1%. She notes that 2 of these subjects had both baseline and Week 180 samples, and there were clear increases in dystrophin in these 2 patients. Of note, Dr. Woodcock points out that although some subjects had Week 180 dystrophin levels similar to the baseline (i.e., close to zero), she would expect this because she would not predict that all individuals would to respond to a drug intervention.

She explains that the issue "…is whether the dystrophin levels found at 180 weeks were within the variability expected for this assay in such patients and, thus, could have arisen by chance, or whether they could have been caused by differences from the controls or from sample

storage as outlined above, or whether they reflected a drug effect, and, thus, whether these data could be seen as adequate and well-controlled."

In the end, taking Dr. Woodcock's arguments into consideration, my view is that the data from Study 202 are *suggestive* of an increase in dystrophin in response to eteplirsen, but the study was not adequate and well controlled. If we accept that there *is* a difference, Study 202 does not reliably speak to the *amount* of dystrophin produced by eteplirsen, given the concerns above. There is only certainty that the largest *possible* amount was 0.93% of normal (on average), and <3% in any individual (if we assume that the quantity was zero at baseline).

Below I will present another concern that leads me to question the veracity of the Western blot data from the Week 180 biopsies from Study 202, based on an issue that Dr. Woodcock did not address in her memo.

c) Study 301:

With the May 26, 2016 goal date approaching, OND and CDER could not reach agreement on the regulatory action for this NDA. In order to gain additional information that might provide evidence of an effect on a surrogate marker that was reasonably likely to predict clinical benefit, we requested that the applicant perform Western blot analyses of skeletal muscle biopsy samples that had been obtained in an ongoing study (Study 301, PROMOVI). These samples were originally planned to be analyzed at the end of the study; however, we requested an interim analyses of a subset of samples. Western blot analyses were performed on paired biceps samples from 13 of the patients. For each of these patients, samples obtained at baseline (prior to treatment) were compared to those obtained at Week 48, after 48 weekly infusions of eteplirsen 30 mg/kg.

The data are shown in Table 1 and the distribution of these changes is shown graphically in Figure 1. Of these 12 patients, 8 (two-thirds) had a change of 0.25% or less; only 1 patient (8%) had a change greater than 1%. The applicant used 3 methods to consider the numerous values below the limit of quantification, but irrespective of the method used, the mean treatment effect was similar, ranging from 0.22% to 0.32% of normal, a change of approximately 2 to 3 parts per thousand that was nevertheless statistically significant.

Table 1: Study 301: Pre- and Post-treatment Values of Becker-Type Dystrophin

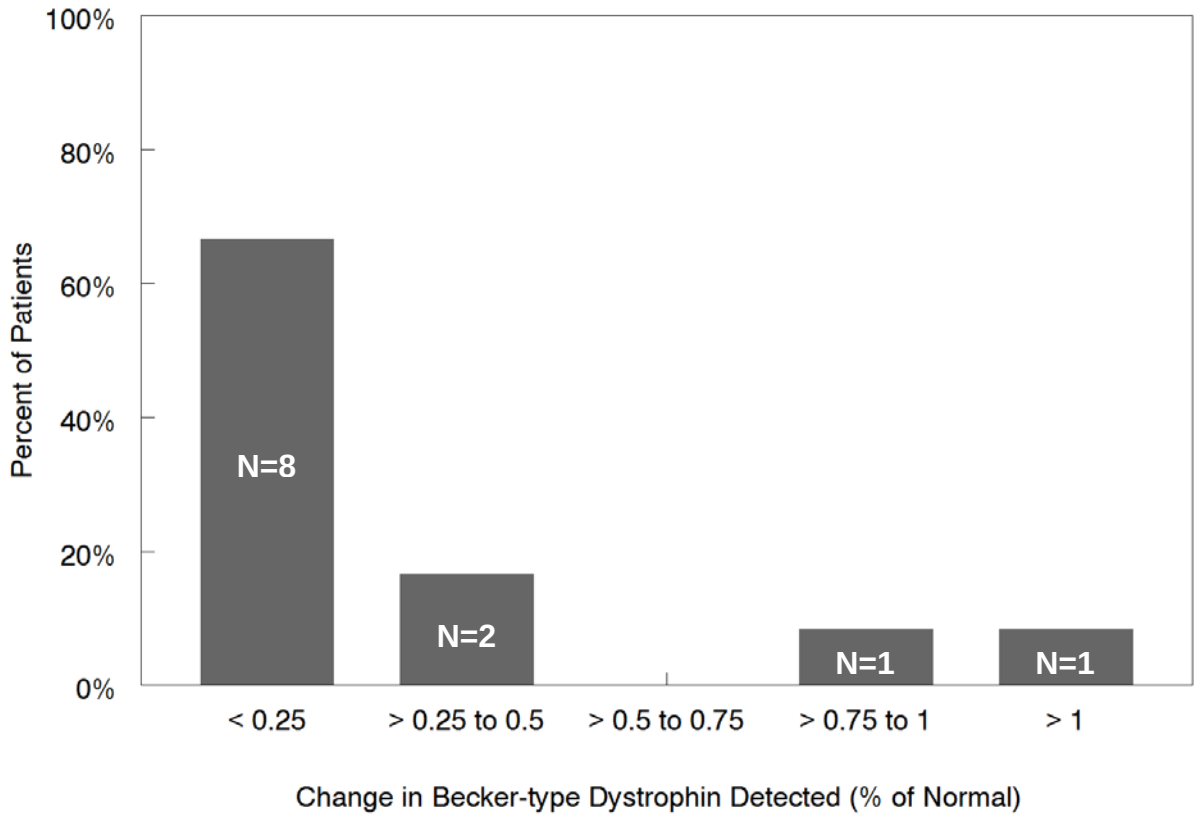
Patient	Time	status	value (%)	mean (%)	delta (%)
1	Baseline	pass	0.15	0.13	**0.13**
		pass	0.11		
	Week 48	pass	0.22	0.26	
		pass	0.29		
2	Baseline	pass	0.35	0.35	**0.01**
		fail	0.26		
	Week 48	pass	0.36	0.36	
		fail	0.12		
3	Baseline	pass	0.06	0.06	**0.31**
		pass	0.06		
	Week 48	pass	0.5	0.37	
		pass	0.24		
4	Baseline	pass	0.04	0.04	**0.06**
		fail	0.06		
	Week 48	pass	0.1	0.1	
		fail	0.19		
5	Baseline	fail	0.1	0.17	**0.85**
		pass	0.17		
	Week 48	fail	0.92	1.02	
		pass	1.02		
6	Baseline	pass	0.37	0.37	**-0.07**
		fail	0.46		
	Week 48	pass	0.3	0.3	
		fail	0.29		
7	Baseline	fail	0.04	0.17	**0.25**
		pass	0.17		
	Week 48	fail	0.22	0.42	
		pass	0.42		

Patient	Time	status	value (%)	mean (%)	delta (%)
8	Baseline	fail	0.08		
		fail	0.14		
	Week 48	fail	0.08		
		fail	0.05		
9	Baseline	fail	0.14	0.24	**1.33**
		pass	0.24		
	Week 48	fail	1.17	1.57	
		pass	1.57		
10	Baseline	pass	0.11	0.11	**0.01**
		fail	0.05		
	Week 48	pass	0.12	0.12	
		fail	0.11		
11	Baseline	pass	0.01	0.05	**0.43**
		pass	0.08		
	Week 48	pass	0.31	0.47	
		pass	0.63		
12	Baseline	pass	0.02	0.02	**0.07**
		fail	0		
	Week 48	pass	0.09	0.09	
		fail	0.01		
13	Baseline	fail	0.34	0.18	**0.03**
		pass	0.18		
	Week 48	fail	0.34	0.21	
		pass	0.21		

All parties agree that these data were obtained from an adequate and well controlled study, and that there is a statistically significant effect of eteplirsen. The disagreement is whether or not the dystrophin production is at a meaningful level that is reasonably likely to predict clinical benefit.

To the extent that one can compare results across studies, these changes in dystrophin are even lower than the values obtained from Study 201/202 (the latter represent the quantity detected at Week 180, not the treatment effect). Dr. Woodcock wrote that "Only 2 of 12 patients achieved a level over 1% of normal control." Her characterization refers to the amount of protein *detected* at Week 48, not the *change* in protein. In fact, only a single patient out of 12 had a *treatment effect* that exceeded 1%.

Figure 1: Study 301: Distribution of Changes in Becker-type Dystrophin in 12 Patients



3. Semi-quantitative assessment of dystrophin in muscle tissue by immunohistochemistry

Study 201/202 – Data through Week 48

Dystrophin production was assessed in Study 201 using immunohistochemistry, a standard laboratory procedure used primarily to localize proteins in tissue sections, but also used as a semi-quantitative method to measure dystrophin levels. Muscle samples were analyzed at baseline, and at Weeks 12, 24, and 48.

Dr. Woodcock notes "A finding of increased dystrophin was also seen in several IHC assays performed by the applicant." She explains that several baseline and other pre-Week 180 assays were performed (from Study 201/202), but the validity of the results was questioned at the FDA inspection because of methodological issues, and so she does not consider these data further.

I do not agree with Dr. Woodcock's outright rejection of these data. In fact, FDA requested a re-reading of the stored images by 3 masked pathologists under improved viewing conditions. We did not request any changes in immunohistochemistry methods or techniques, other than a different approach for selecting microscopic fields for image capture and analysis. Thus, we stressed that their stored images could provide useful data if properly read. The re-read

showed a nominally statistically significant increase in dystrophin in response to eteplirsen for the low dose group, but not the high dose group. (The *p*-value is nominal because the type-I error rate was not controlled for multiplicity.) Moreover, for the 4 patients who had received placebo through Week 24 and then switched to eteplirsen, there was no increase in dystrophin at Week 48.

Study 201/202 – Week 180 Data

The applicant performed immunostaining along with Western blot analyses from the skeletal muscle biopsies obtained at Week 180.

Importantly, prior to performing these analyses, the applicant made changes to the immunohistochemistry protocol with the intent of decreasing non-specific staining. Dr. Woodcock details the technical factors in her memo. Their aim was to determine the treatment effect for each patient, by comparing dystrophin levels at baseline and Week 180. Frozen archived baseline tissue was available for only 3 of the patients, however, and so the applicant supplemented these samples with muscle tissue from 6 untreated external DMD patients, together to be compared to the Week 180 levels. Images were read by the same 3 pathologists, masked to treatment group.

Because external controls were used, the comparison of pre- vs. post-treatment values suffers from the same problems described for the Western blot analyses (i.e., different patients, different muscles, and possible loss of immunoreactive dystrophin with long-term storage).

These concerns notwithstanding, the applicant claimed a remarkable increase in dystrophin immunostaining at Week 180: the 9 baseline samples (from 3 patients in Study 201/202 and 6 external controls) showed 1.1% ± 1.3% positive fibers (mean ± SD), whereas the Week 180 samples (from 11 patients in Study 201/202) showed 17.4% ± 10.0% positive fibers. I will note that FDA made no attempt to inspect or oversee these analyses.

Given that the original analysis showed, at baseline, 13% positive fibers for patients in Study 201/202, it is important to understand why the results from a new immunostaining protocol provided results of 1.1%, an order of magnitude lower.

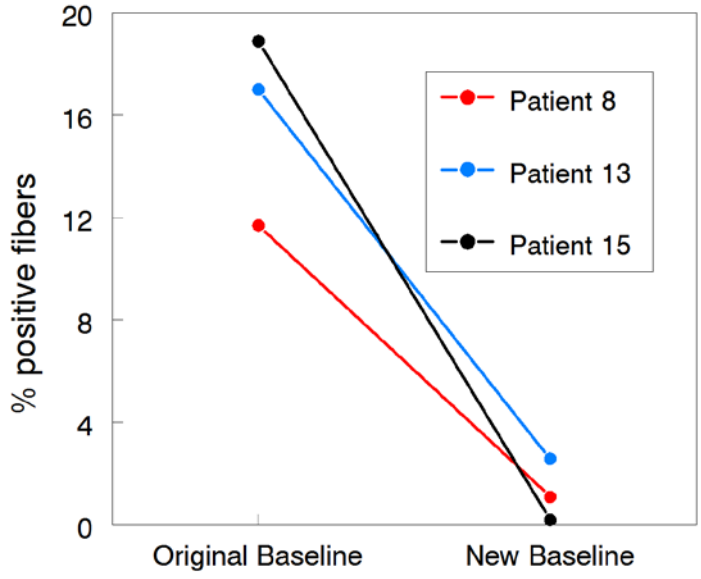
As noted above, there were 3 patients in Study 201/202 with adequate archived tissue from baseline, which permitted a new immunohistochemistry analysis and a comparison of results between the old and new methods. Figure 2 shows how the two methods compare.

These are essentially replicate analyses of a single tissue sample using the two immunohistochemistry methods. There is an inexplicable difference of more than an order of magnitude between results of the old and new immunohistochemistry protocols. Such marked differences raise concerns with respect to the validity of the applicant's methods, and make interpretation impossible.

The disparity also underscores the difficulty of comparing results of immunohistochemical analyses for dystrophin across laboratories, or, for that matter, within the same laboratory.

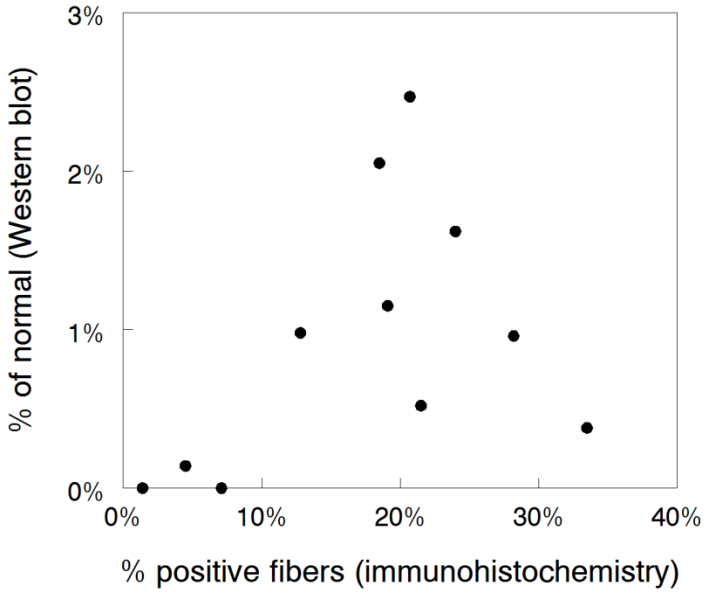
The integrity of the applicant's data is further called into question by lack of agreement between the immunohistochemistry and Western blot methods, i.e., a lack of internal consistency. The applicant claims to have enhanced both the immunohistochemistry methods and the Western blot methods in preparation for processing the Week 180 biopsies. Following these methodological improvements, single tissue blocks were subjected to both analyses – analyses considered to be complementary. Yet the lack of concordance between these two assessments of dystrophin levels is striking (Figure 3).

Figure 2: Comparison of Results from the New and Old Immunohistochemistry Protocols – Lack of Agreement for 3 Patients in Study 201/202



It is simply not possible to determine whether the immunohistochemistry methods are inaccurate, the Western blot methods are inaccurate, or both methods are inaccurate. In light of the discordance between methods, the issues with the control samples, and the order-of-magnitude discrepancy between the old and new immunohistochemistry protocols, these data provide little confidence that the study was designed well enough so as to be able "to distinguish the effect of a drug from other influences, such as spontaneous change…, placebo effect, or biased observation" (§314.126).

Figure 3: Study 201/202 Week 180 Dystrophin Assessment – Lack of Agreement between Immunohistochemistry and Western Blot



A critical point is that results of immunohistochemistry analyses are method-dependent, and results from different laboratories are not directly comparable. Here we see a striking difference between results of different methods within a single laboratory.

Dr. Woodcock concluded "Although the IHC assays provide only semi-quantitative assessments of dystrophin content, they do support an effect of eteplirsen on the proposed surrogate endpoint (an increase of dystrophin production as a result of drug exposure)."

Although this statement does not constitute an important part of her argument in favor of dystrophin production, I do not agree that the immunohistochemistry data show an increase in dystrophin as a result of drug exposure. Given that changes in the immunohistochemistry protocol led to remarkably disparate results, and in light of the lack of correlation between dystrophin results as determined by immunohistochemistry and Western blot, I question the accuracy and interpretability of the Week 180 immunohistochemistry data. Moreover, the results from the properly blinded re-reading of the original data through the first 48 weeks of Study 201/202 are negative. I do agree, however, that the immunohistochemistry images appear to show dystrophin in the proper location, which helps support proof-of-concept.

In summary, I agree that there are data on dystrophin production from one adequate and well controlled study, Study 301, by Western blot. The amount of dystrophin produced and the likelihood of a clinical effect are discussed below.

B. Is the Effect on the Surrogate Endpoint "Reasonably Likely to Predict Clinical Benefit?"

As noted by Dr. Woodcock, "The usual way to address this question would be to rigorously evaluate what is known about the correlation between dystrophin levels in muscle and expression of disease."

Without restating the details of Dr. Woodcock's discussion, I generally agree with her basic summary of the many challenges of interpretation (quoted below). Most of her discussion speaks to the *uncertainties* inherent in correlating dystrophin levels with disease severity. I strongly agree that we lack a sound basis upon which to relate dystrophin levels observed in this development program to observations in the literature.

"1. The clinical classification of disease severity (i.e., phenotype) in the literature appears broad, variable, and somewhat subjective."

> I agree. And importantly, as Dr. Woodcock notes, "the zone of real interest for this discussion, between DMD and intermediate presentations, is not rigorously categorized."

"2. Much of the prior data reporting the relationship of dystrophin protein levels to phenotype have been from immunohistochemistry studies using a variety of techniques and antibodies."

> I will add that the applicant's own data show a striking difference between results of two somewhat different immunohistochemistry protocols conducted at the same laboratory (Figure 2). Thus, it would be treacherous to try to relate various levels of dystrophin, determined by immunohistochemical methods at various laboratories, to a particular clinical course.

"3. Both IHC analyses and WB results are influenced by the anti-dystrophin antibodies used, as well as other experimental conditions"

> Agree. Thus, is not feasible to relate levels of dystrophin determined by older Western blot methods, which lacked, for example, appropriate internal controls, to levels of dystrophin reported in these eteplirsen studies.

"4. The phenotype is significantly influenced by dystrophin isoform quality as well as dystrophin quantity."

> Agree. It is difficult to predict a protein's function from its structure; even small changes in dystrophin structure can be important.

"5. The literature contains various findings on the relationship of dystrophin expression to clinical status, including the low levels of dystrophin protein of interest in this case."

> Agree. There is little consensus on the relationship between dystrophin expression and clinical course at the low levels observed in eteplirsen-treated patients.

I also agree with Dr. Woodcock on the following points, and I paraphrase here:

- Dystrophin levels >10% on Western blot are usually associated with a BMD phenotype. Within the BMD phenotype, the relation between disease severity and protein expression is not clear. Protein quality, rather than quantity, may play a key role in determining phenotype in BMD.

- Patients with DMD are usually found to have undetectable levels of dystrophin, or very low levels. Dr. Woodcock notes that she believes the conventional threshold of <10% protein resulting in DMD was based on immunohistochemistry data. She tries to make a conversion between values observed from immunohistochemistry (~10% points higher on immunohistochemistry than Western blot in DMD) and those observed from Western blot, but I caution that immunohistochemistry results, in particular, are highly method-dependent, as noted above.

- Rarely, dystrophin levels in the 3 to 10% range have been associated with Becker Muscular Dystrophy phenotypes. Dr. Woodcock found no evidence of a threshold value for protein content and expression of a DMD phenotype.

Despite the absence of reliable data, Dr. Woodcock concluded that evidence from Western blot and other experiments shows that protein in the range between undetectable and 10% of normal is likely to be very important for clinical presentation, all other things being equal, i.e., mutation status and non-dystrophin-related factors affecting phenotype.

Because of the lack of reliable evidence, I do not agree that the small increase in dystrophin shown in Study 301 is 'reasonably likely' to predict clinical benefit. This is the central issue in this appeal.

The "reasonably likely" question hinges on whether the protein is functional, and whether the quantity is adequate.

These two uncertainties, protein function and protein quantity, are separate issues that must be considered in series. The function of the Becker-type dystrophin detected in Study 301 cannot be assessed. Nevertheless, the review team has been willing to assume that whatever Becker-type dystrophin is produced would function as well as it does in the Becker form of the disease. Although there can be no certainty on this point, the question of function seems small relative to the uncertainty regarding the adequacy of the quantity of protein, and so function is less germane to the question of "reasonably likely." In short, it is the quantity of Becker-type dystrophin produced that is central to the question of 'reasonably likely,' and central to the approvability of this NDA under accelerated approval.

At the outset, it must be stated that the minimum quantity of Becker-type dystrophin that is reasonably likely to predict clinical benefit in patients with DMD is unknown.

There are two ways to consider the quantity of dystrophin produced: as a binary responder analysis and as a mean response. The former has the advantage of considering the possibility that some patients may respond to the treatment whereas others do not; the latter does not allow for this type of consideration.

The problem with a responder analysis is that there are no data upon which to define a threshold for a 'response.' Various cut-points could be selected, but their selection would be arbitrary, and the particular threshold chosen would have a major influence on the effect size.

Here I provide 3 lines of reasoning to support my view that there is not an adequate basis to believe that the small increase in dystrophin shown in Study 301 is reasonably likely to predict clinical benefit: 1) the treatment effect observed cannot be compared or related to levels of dystrophin measured by other laboratories and reported in various publications; 2) the effect size is inadequate on its face; and 3) no evidence of a clinical effect was demonstrated in the eteplirsen development program, and there is no correlation between dystrophin levels as determined by Western blot and clinical outcome.

1) *The treatment effect observed cannot be compared or related to levels of dystrophin measured by other laboratories and reported in various publications.*

In order to place these small quantities of Becker-type dystrophin into a clinical perspective, many have considered publications from laboratories that attempt to relate particular levels of Becker-type dystrophin protein to clinical course, e.g., maintenance of physical function, age at loss of ambulation. Ideally, as suggested by Dr. Woodcock, there would be reliable data showing that Becker-type dystrophin levels in excess of a particular level are associated with a more benign clinical course.

Realistically however, the use of such a framework would be contingent on the ability to make interpretable cross-laboratory comparisons of dystrophin levels, which would require standardized methods to measure dystrophin levels in muscle specimens. Unfortunately, the methods have differed greatly, and the methods in the literature have lacked critical internal controls such as dilution-series. As stressed above, comparison of dystrophin values across laboratories seems unreliable.

With respect to immunohistochemistry analyses, Figure 2 provides ample basis for concern regarding comparability of results using different methods. Results of separate immunohistochemical analyses of skeletal muscle dystrophin, conducted by the same laboratory on single blocks of tissue, differ by more than an order of magnitude. These results underscore the inherent methodological variability of immunohistochemistry assays, and the futility of attempting to compare dystrophin levels across assays/laboratories.

Even with respect to more recent Western blot methods, reproducibility across laboratories is low. As discussed by Dr. Woodcock, Anthony K *et al* (*Neurology* 2014;83;2062) compared results of Western blot analyses from 6 patients (3 with DMD; 3 with Becker Muscular Dystrophy) across 5 experienced laboratories, and found a high degree of variability. Only one of the 5 laboratories had a coefficient of variation (CV = SD/mean X 100) below 0.3%. The authors found that variability was particularly pronounced with low levels of dystrophin – precisely the area of interest here.

During the applicants' presentation at the April 25, 2016 meeting of the Peripheral and Central Nervous System Drugs Advisory Committee, Dr. Kaye, a pediatric neurologist and interim Chief Executive Officer of Sarepta, could not have been more clear in warning us not to make comparisons between their Western blot results and reported data in the literature:

> "Our validated Western blot method, optimized to detect low levels of dystrophin, is arguably the first dystrophin Western blot to be truly quantitative. This was achieved by use of a 5 point calibration curve on each gel and prespecified loading and exposure limits to avoid signal saturation. Furthermore, samples were randomized, blinded and run in duplicate on separate gels. In contrast, the Western blot methods in the majority of historical publications referenced by FDA were performed using older methodology that is semi-quantitative at best. Given these significant methodological differences, it is inappropriate to compare our data to literature approximations." (*Source*: Official transcript of the meeting; underlining for emphasis.)

In summary, the field has not achieved adequate standardization of methods for dystrophin quantification at the very low levels observed in eteplirsen-treated patients; therefore, it is not valid to compare an increase in Becker-type dystrophin of, at best, 2 to 3%, with dystrophin values cited in the literature for other mutations/patient populations, assessed at other laboratories. *If the applicant's results cannot be compared to results in historical publications, then there is simply no way to determine whether the low dystrophin levels in eteplirsen-treated patients are reasonably likely to predict clinical benefit.*

2) The effect size is inadequate on its face.

If one were to assume that it is possible to make cross-laboratory comparisons of dystrophin levels, the *largest* change reliably demonstrated in Study 301, 1.3%, is an order of magnitude less than the minimum dystrophin levels cited to be important in affecting the course of patients with Becker muscular dystrophy (at least 10%).

Some of the better data come from Van den Bergen et al, who studied the relation between dystrophin levels (quantified by Western blot) and clinical severity in 33 patients with Becker Muscular Dystrophy (*J Neurol Neurosurg Psychiatry* 2014; 85:747). Although the authors did not find a linear relationship between dystrophin levels and disease severity, all 4 of their

patients with dystrophin levels <10% showed poor muscle strength and early symptom onset. As discussed by the review team, DMD experts have proposed that "induction of approximately 10% of normal dystrophin levels sets a minimum level to confer measurable clinical benefit."

Initially, the applicant reported results from immunohistochemistry analyses purportedly demonstrating that eteplirsen caused 50 to 60% positive staining of muscle fibers for dystrophin. This seemingly unprecedented achievement aroused much excitement in the field of DMD research and in the DMD patient community. Upon proper re-analysis, however, the numbers were far lower, and rigorous statistical analyses showed that the changes weren't statistically significant. The Western blot analysis from Study 201/202 showed a mean dystrophin level of only 0.93% (range 0 to 2.5%), but these values are of questionable reliability. Finally, an adequate and well controlled study (Study 301) showed a mean change of 3-tenths of a percent (range 0 to 1.3%). Given that dystrophin is a structural protein, it seems highly unlikely that such changes would translate to a clinical effect.

Here are Dr. Woodcock's assertions on this topic:

> "The broad phenotypic distinctions made in the clinic (e.g., DMD vs IMD vs BMD) are different from the prediction of benefit to an individual patient who has a specific baseline dystrophin level and whose mutation and external factors do not change pre- and post-drug. For example, extending ambulation by six months to a year would not normally move a patient from one to another of these categories, but could be very important to quality of life (e.g., as suggested in the Bello study). This is also true for other functional improvements.
>
> For these reasons, incorporating the analysis of dystrophin content discussed above, I conclude that the biochemical data strongly support the idea that low-level increases in dystrophin production are reasonably likely to predict clinical benefit."

I agree that broad phenotypic distinctions made in the clinic (e.g., Duchenne vs. Intermediate vs Becker Muscular Dystrophy) are different than trying to predict benefit to an individual patient on the basis of a particular change in dystrophin. And I agree that extending ambulation by 6 months to a year (or similar improvements in other functional areas) would be extraordinarily important.

But Dr. Woodcock never provides a rational argument – based on reliable data – to support the concept that "…low-level increases in dystrophin production are reasonably likely to predict clinical benefit." She provides no rationale – no link between a mean increase in dystrophin of 3 parts per thousand and clinical benefit.

3) *No evidence of a clinical effect was demonstrated in the eteplirsen development program, and there is no correlation between dystrophin levels as determined by Western blot and clinical outcome.*

Dr. Woodcock states:

> "Additional support for "reasonably likely" comes from the long-term experience with the drug. The sponsor's comparison of the experience of the treated cohort to natural history data does not reach the level of substantiation required for traditional approval

> based on the clinical data. However, it is highly suggestive of improvement in some parameters, in some patients, over natural history. My conclusion is informed by all the caveats expressed in the reviews about the pitfalls of nonrandomized comparisons. Given that the two exon 52 deletion patients in the study had fairly good long-term results in terms of rate of disease progression, the question arises as to whether exon 52 is a prognostic factor that could have skewed the results."

The review team analyzed the clinical data in great detail, and could not reach the conclusion that there was any reliable evidence of improvement relative to the expected natural history of the disease. Study 201 did not show a treatment effect on its 1° clinical endpoint, change in 6-minute walk distance at Week 24. Study 202 failed on the same endpoint at 48 weeks. The course of these Study 201/202 patients, having received eteplirsen for some 3.5 years, was not distinguishable from external control patients (see my review memorandum for more details).

The Advisory Committee voted (7 to 3 with 3 abstentions) that the clinical results of Study 201/202 did not provide substantial evidence that eteplirsen is effective for the treatment of DMD, and their vote was in the face of extraordinary pressure from patients and patient advocates to vote for approval. Two of the 3 "yes" votes were from patient representatives.

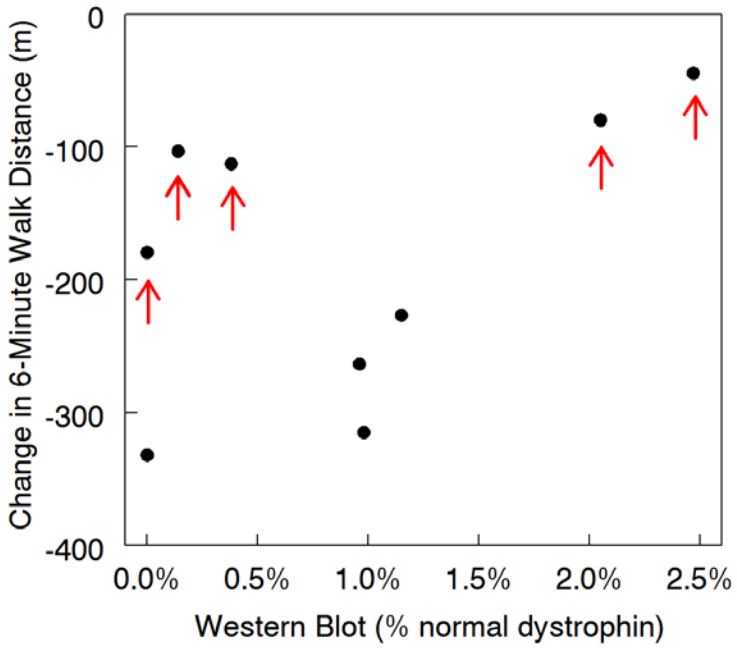
Correlation between dystrophin production and clinical effect

A correlation between dystrophin production (or with less certainty – dystrophin *detected*) and clinical function could provide some support for a conclusion that dystrophin production is reasonably likely to predict clinical benefit.

The applicant collected data on both dystrophin production and physical performance in Study 201/202. On the basis of the data presented in the NDA, the Division concluded that no patient in Study 201/202 clearly deviated from the natural history of the disease. The Division reasoned, therefore, that whatever the quantity of Becker-type dystrophin detected, it did not predict clinical benefit. Thus the Division opined that the clinical data weaken, and do not strengthen, the "reasonably likely" argument.

The Division's view notwithstanding, it is worth considering patients on an individual basis to assess the correlation between the quantity of Becker-type dystrophin detected and changes in physical performance.

Figure 4: Study 201/202 – Lack of Correlation between Quantity of Dystrophin Detected and Preservation of Physical Function (6-Minute Walk Distance)

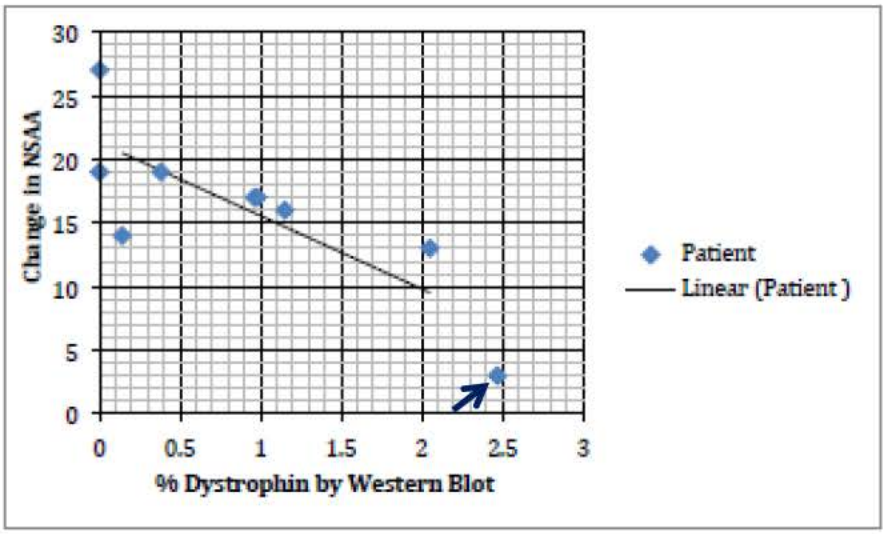


As noted by Dr. Woodcock, the 6-minute walk test results do not show a strong correlation (Figure 4). For the 9 patients in Study 201/202 who remained ambulatory at Week 180 and agreed to undergo a fourth muscle biopsy, the figure shows little correlation between the quantity of dystrophin detected (*x*-axis) and preservation of physical function as assessed by the change in 6-minute walk distance from baseline (*y*-axis) after weekly infusions of eteplirsen for 3 to 3.5 years. For the 5 patients whose 6-minute walk performance was best preserved (red arrows), 2 had the highest dystrophin levels detected in the study (upper right), but 3 had levels that were near-zero (upper left).

Dr. Woodcock also evaluated the North Star Ambulatory Assessment (NSAA) as a function of dystrophin detected in boys who could still walk and who had a dystrophin result at Week 180. She obtained the data from the applicant's briefing document for the Advisory Committee meeting, and found a correlation between dystrophin detected at Week 180 by Western blot and rate of decline in NSAA score through 180 weeks. Her graph is reproduced below:

With respect to the correlation, Dr. Woodcock explained: "This adds additional support to the idea that dystrophin production is "reasonably likely to predict clinical benefit."

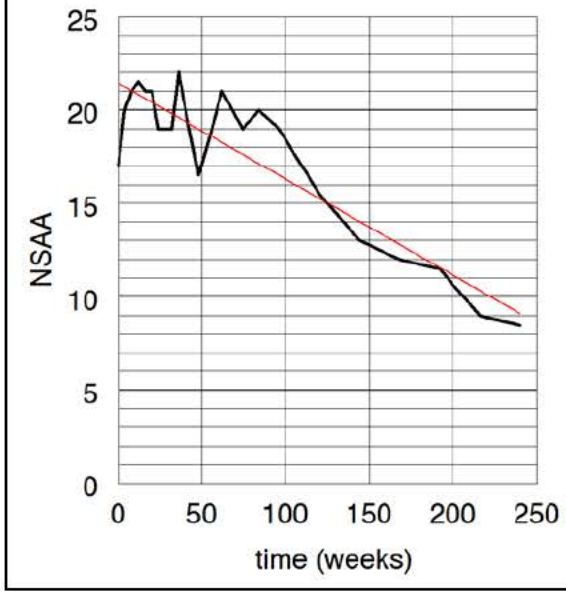
Figure 5: Study 201/202: Analysis of Change in NSAA vs. Expression of Becker-type Dystrophin by Western Blot (Analysis by Dr. Woodcock)



Given that the correlation was driven by the patient depicted at the lower right (blue arrow; dystrophin level = ~2.5%; change in NSAA = 3), I considered the NSAA data from that patient (Figure 6). I found that his course was less benign than would be inferred from a change in NSAA of only 3 units. Specifically, using linear regression (red line in Figure 6), his NSAA score has, instead, worsened by a mean of 2.7 units per year.

I reasoned that inclusion of <u>all</u> of the NSAA data for each patient would provide a more reliable representation of their course than calculating the change between single pre-treatment and post-treatment data points, because of the test-to-test variability (e.g., short-term swings of 4 to 5 points for patient 006). Thus, using linear regression, I calculated the slope of the relationship between NSAA and time for each patient (as per the red line in Figure 6) and plotted the slopes as a function of the dystrophin detected at Week 180. (Slopes were calculated as loss of NSAA units per year.)

Figure 6: Patient 006: NSAA vs Time

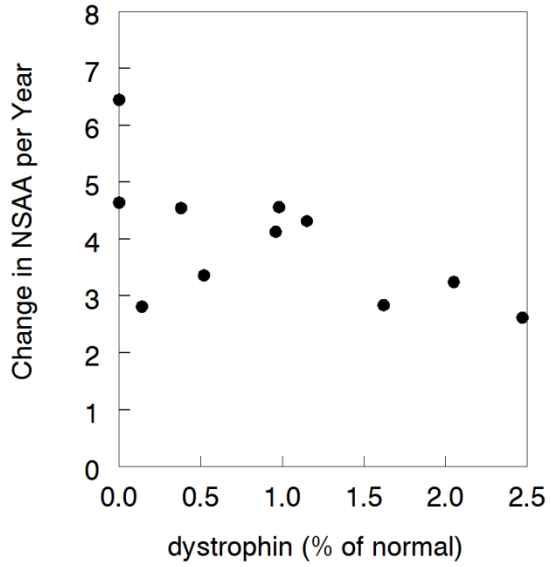


Using this method, there was no correlation (R^2 = 0.36), Figure 7. Importantly, the slight trend apparent here is driven by one or two data points.

Summary:

In summary, I find no evidence that the increase in dystrophin demonstrated in Study 301 is reasonably likely to predict clinical benefit (mean 0.3%, range 0 to 1.3%). The levels of dystrophin linked to various Becker Muscular Dystrophy phenotypes in publications are largely not comparable to dystrophin levels measured in this development program. The applicant's interim CEO correctly urged us not to compare data from their Western blot analyses to historical approximations from the literature. And extremely low levels of dystrophin, as found here, seem particularly difficult to quantify and compare across laboratories. Nevertheless, to the degree that findings can be compared across studies, dystrophin levels of 10% of more would need to be achieved to impact the clinical course. The finding in Study 301 is an order of magnitude below this level.

Figure 7: Study 201/202: Analysis of Change in NSAA (Linear Regression) vs. Expression of Becker-type Dystrophin by Western Blot (My Analysis)



Based on protein levels in other deficiency diseases, the effect size here appears to be too small to provide benefit. If dystrophin were an enzyme that catalyzed a biochemical reaction in myocytes, one might posit that a very small quantity could produce a substantial proportion of the minimum necessary reaction product, and that the increase over baseline might be important because levels are so low in untreated patients. But given that dystrophin is a structural support protein that helps prevent myocyte injury due to stress and strain, I find it difficult to conceive how a treatment effect of 3 parts per thousand could confer clinical benefit. If there were 10 inches of snow on a sidewalk that needed to be cleared, 3 parts per thousand would amount to 1/32nd of an inch. We must also recognize that a treatment that increases dystrophin by 0.3% would seemingly have far less impact than being born with 0.3% more dystrophin, and even *that* seems unlikely to matter.

I can find no precedent of an accelerated approval for a marketing application where the effect size on the surrogate endpoint is as small as 0.3%.

Dr. Woodcock concludes:

> "…my conclusion to rely on the surrogate endpoint described above represents the greatest flexibility possible for FDA while remaining within its statutory framework. In this case, the flexibility is warranted because of several specific factors, including: the life-threatening nature of the disease; the lack of available therapy; the fact that the intended population is a small subset of an already rare disease; and the fact that this is a fatal disease in children. Of note, the therapy has been relatively safe in the clinic, although intravenous administration always carries risk.…Therefore, I find that the

probable benefits outweigh the foreseeable risks and that this application should be approved under 21 CFR 314.510."

As noted in 506(f)(1), the amendments made by the Food and Drug Administration Safety and Innovation Act (FDASIA) "…are intended to encourage the Secretary to utilize innovative and flexible approaches to the assessment of products under accelerated approval for treatments for patients with serious or life-threatening diseases or conditions and unmet medical needs."

Some have interpreted this "flexibility" as a lower standard for demonstration of effectiveness, but this is not true.

Section 506(f)(2) of the FD&C Act specifically notes that drugs granted accelerated approval must meet the same statutory standards for safety and effectiveness as those granted traditional approval, notably the substantial evidence standard of section 505(d) with respect to the drug's claimed effect on a surrogate or intermediate endpoint. These facts have not been altered by FDASIA.

To be clear, 506(f)(2) states: "Nothing in this section shall be construed to alter the standards of evidence under subsection (c) or (d) of section 505 (including the substantial evidence standard in section 505(d)) of this Act or under section 351(a) of the Public Health Service Act. Such sections and standards of evidence apply to the review and approval of products under this section, including whether a product is safe and effective. Nothing in this section alters the ability of the Secretary to rely on evidence that does not come from adequate and well controlled investigations for the purpose of determining whether an endpoint is reasonably likely to predict clinical benefit as described in subsection (b)(1)(B)."

I believe the burden is on Dr. Woodcock to show or explain why production of a near-zero quantity of dystrophin (0.3%) is reasonably likely to predict clinical benefit, and I do not believe her July 14, 2016 memo makes this case. I believe that the available evidence leaves open the possibility that some patients *could* benefit from a small increase in dystrophin, but this possibility does not reach the threshold of being reasonably likely to predict a clinical benefit.

Finally, there was no clinical benefit demonstrated in the development program, and the correlation between dystrophin and clinical effect was poor – not surprising given that the applicant provided analyzable data from only 11 patients.

3. Assessment of Possible Impact to Public Health Should My Position Not be Adopted

The approval of this NDA in its present form would have far reaching negative consequences for the public health.

1. Eteplirsen's risks are certain, whereas its efficacy is not. Having considered Dr. Woodcock's line of reasoning and her desire to approve eteplirsen, the position of the review team in the Division of Neurology Products, the Office of Biometrics, the Office of Clinical Pharmacology, the Office of Drug Evaluation-I, and the Office of New Drugs (verbal acknowledgement from Dr. John Jenkins) is that the applicant has not provided evidence that this drug is effective at the dose studied.

Dr. Woodcock notes that "…the therapy has been relatively safe in the clinic."

The reality is that only a few dozen patients have been exposed to the drug, such that the safety profile is not well characterized. A closely related drug being studied under a (b) (4) With additional experience, important toxicity may emerge for eteplirsen. It is known that many patients in these studies are now receiving infusions through indwelling catheters. Maintenance of vascular access in patients on chronic corticosteroids poses a certain risk of infections. Although we are not yet aware of any infection-related adverse reactions, there would definitely be serious infections and possibly deaths if this drug is marketed, yet evidence of efficacy is lacking.

2. By allowing the marketing of an ineffective drug, essentially a scientifically elegant placebo, thousands of patients and their families would be given false hope in exchange for hardship and risk. I argue that this would be unethical and counterproductive. There could also be significant and unjustified financial costs – if not to patients, to society.

   The prospect of providing false hope to desperate patients from a promising but ineffective therapy recalls the experience with *transmyocardial laser revascularization* (TMLR). In the 1990s, patients with coronary atherosclerosis and severe angina who were poor candidates for conventional revascularization procedures ("no-option" patients) underwent a thoracotomy (opening of the chest cavity) to enable use of a laser to create channels through the heart muscle. Ostensibly, these channels provided conduits for blood to flow from inside the left ventricle to the myocardium. Conduct of sham-controlled studies was impossible; studies were essentially baseline-controlled or historically-controlled. Large treatment effects were reported by a number of investigators, generally from small studies. There were marked increases in treadmill exercise time and relief of angina, with effects sustained for more than a year in some cases. Although many in the cardiology community raised concerns about expectation bias and were highly skeptical of the results, to some the effects seemed larger and more durable than could possibly be explained by expectation bias, i.e., a placebo effect. Thousands of patients underwent this invasive procedure with the hope of angina relief. Some years later, with improvements in technology, the conduct of sham-controlled studies became feasible, and TMLR was not found to be effective. The false hope was ultimately dispelled with the publication of two Cochrane Reviews.[3] These reviews found the appearance of a marked treatment effect, but 30-day mortality was 6.8% in the TMLR group vs. 0.8% in the no-treatment group. They noted "The assessment of subjective outcomes, such as improvement in angina, was affected by a high risk of bias and this may explain the differences found." In this case, the cost of false hope was ~6% mortality in the first 30 days post-op.

   I will also note that the primary endpoint of these laser studies was generally exercise capacity – the same type of endpoint used in the eteplirsen DMD development program, also for "no option" patients.

3. The accelerated approval pathway is designed to expedite the availability of promising new therapies to patients with serious conditions, especially when there are no satisfactory

[3] Cochrane Database of Systematic Reviews 2015, Issue 2. Art. No.: CD003712. DOI: 10.1002/14651858.CD003712.pub3

alternative therapies, while preserving standards for safety and effectiveness. For drugs granted accelerated approval, postmarketing confirmatory trials are required to verify and describe the anticipated clinical benefit, and FDA may withdraw approval of a drug if a trial required for verification of the predicted clinical benefit fails.

In reality, it is difficult to withdrawal a drug that is deemed to be effective, or possibly effective, by patients with severe diseases and limited treatment options. FDA has not succeeded in withdrawing the marketing of a single drug for lack of verification of clinical benefit following accelerated approval. The reality is that if eteplirsen is given accelerated approval, it is highly likely to remain on the market indefinitely, irrespective of whether or not efficacy is verified.

4. With the false perception that eteplirsen is effective, patients who are gaining benefit from steroids but experiencing untoward side effects might be inclined to taper or stop them, which could lead to more rapid disease progression.

5. False scientific conclusions have the potential to mislead the field of medicine, slowing progress in finding and developing therapies that actually *are* effective. For example, consider the scenario of a related drug with far greater potential to promote dystrophin production in patients with DMD. In order for a sponsor to study such a drug, patients would likely have to agree to discontinue eteplirsen, and few patients may be willing to do so. In short, approval of an ineffective therapy has the potential to discourage or inhibit the development of other drugs that *are* effective, and this impact can be significant.

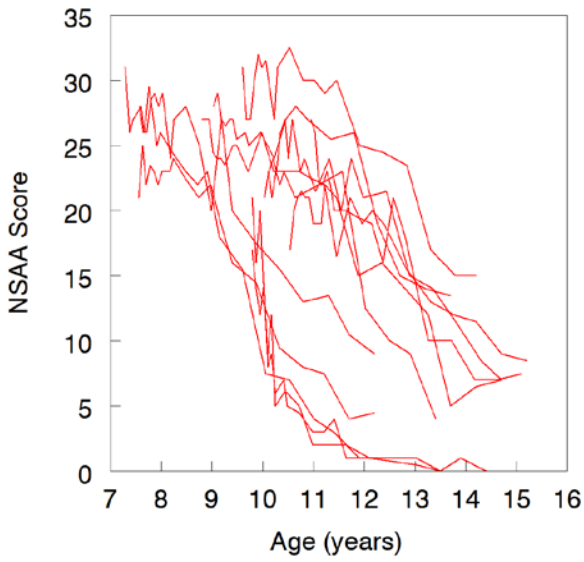
6. Accelerated approval would lower the evidentiary standard for effectiveness to an unprecedented nadir. The amount of dystrophin produced in Study 301 is so meager that it could be considered to be tantamount to <u>any</u> increase in dystrophin. In other words, if a statistically significant change of 0.3% – a mere 3 parts out of a thousand – is considered adequate to support accelerated approval here, then the question arises as to whether there would be <u>any</u> statistically significant change that would be too small to be considered "reasonably likely" to support accelerated approval. Similarly, if a 'responder' had been defined as a patient with an increase in dystrophin of ≥1% (and there is no basis to accept such a low threshold), there would have been only a single responder in Study 301. If we were to adopt the concept that, for rare diseases, accelerated approval could be supported by <u>any</u> statistically significant change in an appropriate surrogate, or a response in a <u>single</u> patient, we would enable accelerated approval of a myriad of drugs for rare diseases. No doubt there are some who would applaud this as an advance. But a standard this low would undercut FDA's ability to ensure that drugs that are approved are effective; it would call into question much of what we do. Lowering the bar to this level would be tantamount to rolling back the 1962 Kefauver-Harris Drug Amendments to the Federal Food, Drug and Cosmetic (FD&C) Act, which have served Americans well for some 54 years.

7. With accelerated approval of this NDA, there would be highly detrimental effects on drug development. Traditional drug development for rare diseases might be replaced by a system where small, baseline-controlled, proof-of-concept studies designed to show <u>any</u> change in a surrogate marker would provide a basis for accelerated approval, assuming that the pathogenesis of the disease was well understood and that the surrogate was directly on the causal path. There would be little reason to pursue adequately controlled clinical trials to support efficacy prior to accelerated approval; in fact, the possibility of

failure would provide a disincentive to conduct such trials. For example, a gene therapy designed to produce a missing clotting factor could receive accelerated approval on the basis of a tiny yet inconsequential change in levels of the factor, or a more robust response in a single patient. In short, the precedent set here could lead to the approval of drugs for rare diseases without substantial evidence of effectiveness.

8. Even if the 30 mg/kg/week dose were considered to have a meaningful effect on the surrogate endpoint, we already know this dose is sub-therapeutic. We know this because patients who have been receiving this eteplirsen dose for some 3.5 years have been progressing at a rate that is similar to that expected, based on the natural history of the disease (Figure 8). I question the ethics of approving or prescribing a drug for a fatal disease at a dose that is very likely to be sub-therapeutic, when the consequence of a sub-therapeutic dose is clinical deterioration and death. The figure shows the unremitting progression in the patients in Study 201/202, based on changes in NSAA.

Figure 8: Study 201/202 – Individual NSAA Performance by Age



9. Approval of this NDA would send the signal that political pressure and even intimidation – not science – guides FDA decisions, with extremely negative consequences (See Grainger D., 11/30/15. "DMD Drugs: an existential threat to FDA," *Forbes*[4]). The public is well aware of this development program: the meager size of the study population, the marginal (at best) effect size, the Division's dim view of the efficacy data, and the robust activism of some members of the DMD community. Many would be amazed at an approval action, because other DMD drugs, recently turned down for approval, appeared to provide stronger evidence of efficacy.

   FDA and Congress were bombarded with correspondence – pleas urging approval of this NDA. More than 50 speakers registered to speak at the April Advisory Committee meeting. I received 2,792 emails urging approval. Here is an example of the body of an email I received last week:

   > "Dear Dr. califf: How is it that everyone in and around DMD understands this simple Idea and the science geniuses at FDA don't? You stupid f_ _ _ers are costing each and every DMD kids days of their lives with your Moronic Dystrophin dance. Time to get a

[4] downloaded 7/18/16 at http://www.forbes.com/sites/davidgrainger/2015/11/30/dmd-drugs-an-existential-threat-to-the-fda/#5ffc712455f7

f_ _ _ing clue [redacted (b) (6)] ”

The ramifications here are profound. The public will perceive that it was their unprecedented lobbying efforts that made the difference and earned eteplirsen its accelerated approval. For the future, this will have the effect of strongly encouraging public activism and intimidation as a substitute for data, which is one of the worse possible consequences for communities with rare diseases. This type of activism is not what was envisioned for patient-focused drug development.

4. Detailed Description of the History of the Dispute, Including My Description of the Center SDR Procedures Followed and/or Not Followed, Dates of Meetings, and Decisions Rendered Throughout the Process

The following table shows the dates and main activities for 15 Center Director Briefings associated with the development of this drug: 8 Center Director Briefings took place during the IND phase of development, *prior to submission of the NDA*, and 7 Center Director Briefings took place during review of this NDA.

DATE	MEETING	DETAILS
7/17/2013	Center Director Briefing	Follow up on Action Item from 3/13/13 EOP2 Meeting: Sarepta has submitted a comprehensive discussion of the issues from the EOP2 mtg. To discuss the suitability to file the NDA for Subpart H approval.
10/18/2013	Center Director Briefing	Dr. Unger presented an overview and Dr. Farkas had a slide presentation on drisapersen and eteplirsen data. Discussion: 1. Plan to have a manufacturing facility visit by ONDQA - to observe process and obtain yield calculation. Sponsor is expecting to have 2nd batch in Dec 2013. Determine how much product the sponsor has. 2. OBP: recommended to establish specificity of the antibody and variability of the assay. 3. Next trial - plan to have OSI group to observe the conduct. 4. Need data from the [redacted (b) (4)]) trial. DNP has previously requested the [redacted (b) (4)] data from [redacted (b) (4)] but did not get any response. Dr. Woodcock will initiate an inquiry to the sponsor (raw data). 5. The Agency needs to assist Sarepta (characterize biomarker, CMC facility, observe 6MWT, etc.) 6. 2nd Internal Meeting (Drs. Woodcock, Temple, Jenkins, Unger and Neuro) before the 11/8/13 sponsor meeting. Discuss further what to convey to Sarepta.
10/28/2013	Center Director Briefing (continuation of 10/18/13 meeting)	Suggestions/Recommendations for DNP to Consider: -- We have concluded that we will not ask for biopsy until (we understand the histopathology and are) we're certain what is a quantitative measure and identified the surrogate marker for the study. -- Tell the sponsor that we have changed our view for the quantitative measure of truncated dystrophin as a surrogate PD marker used in their study, because of the recent natural history

		study and (b) (4) . -- Dr. Woodcock wants to have a comprehensive literature review to fully understand what's this mean of the deletions, mutations, or duplications in the dystrophin gene, or this exon 51 of dystrophin mRNA ((Office of Translational Science)I believe this task was assigned to a different group). -- To ask the Sponsor to provide their production schedule. I believe Dr. Woodcock wants to understand the amount of production and determine if the company can provide the drugs to those DMD patients in the future. -- To suggest that the Sponsor consider enrolling patients younger in age (like starting with 5yrs) in their clinical study. -- To ask the Sponsor if they could provide drugs for compassionate use to patients (who are very sick or those were in the drisapersen trial previously). -- Schedule a T-con with (b) (4) to discuss (b) (4) data
1/17/2014	Center Director Briefing	Request: Team to present DMD drugs study design to Dr. Woodcock – Path forward for Sarepta (& (b) (4))
2/6/2014	Center Director Briefing	DMD drugs study design (Discuss Sarepta path forward) Action items: (a) Request biomarker data from the sponsor - done TC on 2/7/14(b) If data interpretable, meet with sponsor for a brainstorming session. Then follow-up with Advice Letter
3/5/2014	Center Director Briefing	Dr. Ash Rao presented biomarker data findings (including Drs. Woodcock, Jenkins, Temple, Unger, Moscicki) Team discussed path forward. Action Item: to invite Sarepta for a brainstorming discussion.
3/19/2014	Sponsor Meeting, with Center Director	brainstorming discussion - study design and path forward Action: Sarepta to submit proposed studies and next steps
4/2/2014	Center Director Briefing	Drs. Woodcock, Moscicki, Temple, Unger Discuss proposal & comments to sponsor ~Advice Letter-include previous meeting discussions ~FDA workshop – biomarker ~Work w/ sponsor on dystrophin biomarker ~Natural history raw data - primary investigators
6/26/2015		**SUBMISSION OF NDA**
12/9/2015	**Center Director Briefing**	To brief on the current status of eteplirsen review in advance of the planned Jan 22, 2016 AC meeting. To discuss the application and the plan of action.
1/13/2016	**Center Director Briefing**	To review the slide presentation and plan of action for eteplirsen, that will be presented during the Advisory Committee Meeting on January 22, 2016 to senior leadership.
2/10/2016	**Center Director Briefing**	To discuss the ongoing review of the NDA, and what will be presented during the Advisory Committee Meeting in April. To discuss the strengths, limitations, and uncertainties of the data, particularly with respect to the comparison between the open-label eteplirsen group and a contemporary untreated external control group.
4/15/2016	**Center Director Briefing**	To discuss the statistical review of the CINRG data. To discuss the review of data on DMD that was conducted by the Cooperative International Neuromuscular Research Group
4/25/2016	Advisory Committee Meeting	
5/4/2016	**Center Director Briefing**	Discuss the outcome and plan of actions for the application post advisory committee meeting

5/31/2016	**Center Director Briefing**	Discuss reviews conducted by the review team and leadership along with any additional information obtained from the sponsor. Discussed Dr. Woodcock's memo. Timeline for reviews due to Dr. Woodcock.
7/6/2016	**Center Director Briefing**	1. The levels of dystrophin observed in 12 DMD patients from the recent interim analysis of an ongoing trial and whether the levels seen can be interpreted to be "reasonably likely to predict clinical benefit" and used as a surrogate endpoint to support accelerated approval. 2. The design of one or more PMR trials to confirm clinical benefit of eteplirsen if it is approved under accelerated approval. 3. Description of the available clinical data in the drug label if approved.

Based on my years of experience in Office of Drug Evaluation-I, the Center Director's direct involvement with this drug, compared to other development programs, has been unprecedented. She also attended the April meeting of the Peripheral and Central Nervous System Drugs Advisory Committee, where she spoke and interjected a number of important comments.

There is no question that there has been adequate time and place for the discussion of various views. I will note, however, that I found it unfortunate that the Center Director made clear her intent to approve the drug at a briefing with the review team on May 4, 2016, before she had seen drafts of the Division's final review memorandum or my review memorandum. Prior to reading our reviews, Dr. Woodcock stated that she had already "…reached a different conclusion…." than the review team.

5. Action, Decision or Remedy Sought

Although the above paragraph could be considered grounds for an appeal based on process, I seek instead a scientific review on the matter of whether or not there is substantial evidence of a quantitative effect on dystrophin protein that is reasonably likely to predict clinical benefit. I maintain, along with the Division of Neurology Products, Office of Biometrics, Office of Clinical Pharmacology, Office of New Drugs, and the majority of the members of the Peripheral and Central Nervous System Drugs Advisory Committee, that substantial evidence is lacking to support either a conventional or accelerated approval, and that a *complete response* should be issued for this NDA.

The unprecedented finding of an increase in dystrophin protein in response to eteplirsen establishes proof-of-concept and provides great promise that this drug, or other therapies, will eventually be capable of ameliorating the fundamental genetic defect of DMD, but the effect size here is insufficient at the tested doses.

6. Path Forward

Based on the quantity of Becker-type dystrophin produced in Study 301 and the clinical findings in Study 201/202, additional studies at this dose are unlikely to support any type of approval, i.e., the data obtained for eteplirsen at doses of 30 and 50 mg/kg/week are fairly solid, but they do not support efficacy.

I remain comfortable with the concept that substantial evidence of dystrophin production from adequate and well controlled trials could support accelerated approval, but it is clear that higher doses are needed, and greater quantities of dystrophin would need to be produced. The path to a conventional approval would require a double-blind, placebo-controlled (or multi-dose) study, at least one year in duration, using some measure of physical performance as the primary endpoint, again, testing higher doses.

The applicant is continuing to enroll Study 301 (PROMOVI), an open-label, multicenter, 48-week study in patients with DMD amenable to skipping exon 51. All patients are receiving eteplirsen, 30 mg/kg/week as an IV infusion.

My suggestion for a path to approval is to randomize patients in the ongoing Study 301 to:

1) either remain on 30 mg/kg/week; or
2) have their dose significantly increased. This could be done through use of a higher dose, through more frequent dosing intervals (with dummy infusions), or both. Given that many patients receive eteplirsen through indwelling IV lines and no significant infusion reactions have occurred, perhaps these infusions could be performed at home. For example, the study could compare 30 mg/kg weekly to 30 mg/kg daily. Patients who do not tolerate more frequent dosing could have their doses decreased, as needed. Based on non-clinical findings, monitoring would need to be in place to assess renal toxicity.

Patients and investigators would be blinded to treatment group. For accelerated approval, the primary endpoint would be dystrophin production, comparing the higher and lower doses. For standard approval, the primary endpoint would be a test(s) of physical performance such as NSAA or rise time.

Such a trial would be methodologically sound and ethical. Virtually everyone, patients and physicians alike, would want to know whether higher eteplirsen doses would increase dystrophin production, and would have equipoise for participation. Although there is concern regarding performance of muscle biopsies in patients randomized to placebo, this would not be a concern here with all patients receiving active drug. And I would recommend that the applicant forego immunohistochemistry studies in favor of Western blot analyses, such that needle biopsies with local anesthesia would be sufficient (rather than open biopsies with more intensive anesthesia and greater morbidity).

I also believe that it would be desirable for the company to provide access to eteplirsen for DMD patients through expanded access programs, with cost recovery, while an adequate dose-finding study is conducted.

FDA is charged with the responsibility of ensuring that drugs are shown to be effective prior to marketing, based on substantial evidence. If we were to approve eteplirsen without substantial evidence of effectiveness, or on the basis of a surrogate endpoint with a trivial treatment effect, we would quickly find ourselves in the position of having to approve a myriad of ineffective treatments for groups of desperate patients, in essence, allowing marketing based on desperation, patient lobbying, and the desire and need of hope. If we were to turn the clock back to the days prior to the 1962 Kefauver Harris Amendments to the Federal Food, Drug, and Cosmetic Act, the damage to society and the field of evidentiary medicine would be enormous.