

Office of Drug Evaluation-I: Decisional Memo

Date	July 15, 2016
From	Ellis F. Unger, MD, Director Office of Drug Evaluation-I, Office of New Drugs, CDER
Subject	Office Director Decisional Memo
New Drug Application (NDA) #	206488
Applicant Name	Sarepta Therapeutics
Date of Submission	June 26, 2015
PDUFA Goal Date	May 26, 2016 (post-3-month extension for major amendment)
Proprietary Name/ Established (USAN) Name	EXONDYS 51™ eteplirsen injection
Dosage Forms/ Strengths	2 mL single-use vials containing 100 mg (50 mg/mL) eteplirsen 10 mL single-use vials containing 500 mg (50 mg/mL) eteplirsen
Indication originally sought by applicant (see page 29 for final)	“EXONDYS 51 is indicated for the treatment of Duchenne muscular dystrophy (DMD) in patients who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping. This indication is approved on an intermediate endpoint demonstrating delayed disease progression as measured by the 6 minute walk test [see Clinical Studies (14)]. Continued benefit will be evaluated through confirmatory trials.”
Action:	<i>Complete response</i>

Material Reviewed/Consulted - Action Package, including:	
Project Manager	Yuet (Fannie) Choy, Laurie Kelley
Medical Officer/Clinical	Christopher Breder
Clinical Pharmacology/Pharmacometrics	Atul Bhattaram, Ta-Chen Wu, Hobart Rogers, Kevin Krudys, Angela Men, Christian Grimstein, Mehul Mehta
Statistical Review	Xiang Ling, Kun Jin, Hsien Ming (Jim) Hung
Pharmacology Toxicology	David Hawver, Lois Freed, Paul Brown
Office of Biotechnology Products	Ashutosh Rao, Amy Rosenberg
Office of New Drug Quality Assessment	Joseph Leginus, Mari Chelliah, Sung Kim, Denise Miller, Zhong Li, Dahlia Woody, Martha Heimann, James Laurenson, Donna Christner, Neal Sweeney, Edwin Jao, Zhihao Peter Qiu, Wendy Wilson-Lee, M. Scott Furness
Office of Scientific Investigations	Antoine El Hage, Cara Alfaro, Susan Thompson, Kassa Ayalew, Ni Aye Khin
Method Validation	Michael Hadwiger, Michael Trehy
Statistical Review – Stability data	Zhuang Miao, Xiaoyu Dong, Meiyu Shen, Yi Tsong
Office of Prescription Drug Promotion	Aline Moukhtara
Division of Medication Error Prevention and Analysis	Deborah Meyers, Justine Harris, Danielle Harris, Todd Bridges
Division of Risk Management	Robert Pratt, Jamie Parker, Kellie Taylor, Cynthia LaCivita
Associate Director for Labeling	Tracy Peters
Cross-Discipline Team Leader	Ronald Farkas
Deputy Director, Division of Neurology Products	Eric Bastings

1. Introduction

Sarepta Therapeutics is seeking accelerated approval for eteplirsen for the proposed indication:

“EXONDYS 51 is indicated for the treatment of Duchenne muscular dystrophy (DMD) in patients who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping. This indication is approved on an intermediate endpoint demonstrating delayed disease progression as measured by the 6 minute walk test [see Clinical Studies (14)]. Continued benefit will be evaluated through confirmatory trials.”

I agree with the views of the Division of Neurology Products (DNP), the Office of Biometrics, and the Office of Clinical Pharmacology that the applicant has not provided substantial evidence of effectiveness from adequate and well controlled trials to support conventional approval. I also agree that the applicant has not provided support for accelerated approval, i.e., evidence from adequate and well controlled trials of an effect on a biomarker that is reasonably likely to predict effectiveness. Thus, I agree with the DNP recommendation to issue a *Complete Response* for this application.

2. Background

Description:

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, restoration of the mRNA reading frame would permit translation of an internally truncated, but nevertheless functional form of the dystrophin protein. The drug is targeted specifically for patients with DMD “who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping.” It is not clear which of the specific mutations are amenable to exon 51 skipping.

PMOs are a class of synthetic molecules based upon a redesign of the natural nucleic acid structure. They are distinguished from native DNA and RNA because of a 6-membered morpholino ring that replaces the 5-membered ring found in native DNA and RNA. Each morpholino ring is linked through an uncharged phosphorodiamidate moiety rather than the negatively charged phosphate linkage that is present in native DNA and RNA. Each morpholino subunit contains one of the heterocyclic bases found in DNA (adenine, cytosine, guanine, or thymine). Eteplirsen contains 30 linked subunits. The molecular formula of eteplirsen is $C_{364}H_{569}N_{177}O_{122}P_{30}$ and the molecular weight is 10.3 kilodaltons.

Disease Background:

Duchenne muscular dystrophy is an X-linked recessive neuromuscular disorder caused by mutations of the dystrophin gene located on the short arm of the X chromosome. These mutations disrupt the mRNA reading frame, leading to the absence or near-absence of dystrophin protein in muscle cells. The disorder affects 1 in ~3,600 boys (~1 in 10,000 to 14,000 males). Patients who are amenable to skipping exon 51 constitute ~13% of the DMD patient population.

Dystrophin is thought to maintain the structural integrity of the muscle cell membrane by connecting the cytoskeleton to the underlying extracellular matrix, and acting as a scaffold for several molecules that also contribute to normal muscle physiology. Absence of dystrophin leads to mitochondrial dysfunction and damage, with inflammatory processes also appearing to contribute to muscle pathology. Muscle fibers ultimately undergo necrosis with replacement by adipose and connective tissue. Principal disease manifestations include progressive degeneration of skeletal and cardiac muscle, leading to 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. Currently, glucocorticoid therapy is the cornerstone of clinical management, and is widely believed to delay loss of ambulation and respiratory decline by several years. Ventilatory assistance and physiotherapy are also thought to improve survival for DMD patients.

3. Product Quality

From a product quality perspective, NDA 206488 is recommended for approval. Eteplirsen would be marketed as a sterile, aqueous, preservative-free, concentrated solution for dilution prior to IV administration, to be supplied in single-use glass vials containing 100 mg or 500 mg eteplirsen (50 mg/mL).

OPQ recommends the following post-marketing commitments (PMCs), to be fulfilled no later than one year following NDA approval:

1. Investigate the root cause of the increasing assay trend observed in the drug product stability study.
2. Revalidate the accuracy of the in-process (b) (4) method used during drug product manufacture.
3. Revalidate the robustness of the in-process (b) (4) method in terms of (b) (4) (b) (4).
4. Investigate the consistent bias in the in-process (b) (4) results and the release (b) (4) results.

4. Nonclinical Pharmacology/Toxicology:

From a nonclinical perspective, NDA 206488 is recommended for approval. Pivotal toxicology studies were conducted in male monkeys (39-week study) and juvenile male rats (10-week study). A 26-week study was conducted in male transgenic mdx mice using a mouse-specific surrogate (AVI-4225). In all 3 species, the kidney was identified as the 1° target organ, with dose-dependent renal tubular cytoplasmic basophilia and/or vacuolation and tubular degeneration and necrosis, primarily at the highest doses tested.

Dilatation of the lateral ventricles of the brain was observed at mid and high doses in the *mdx* mouse study. The mechanism of this effect and its relevance to humans are unknown. In juvenile rats, slight reductions in bone length, width, area, mineral content, and mineral density were observed at the high dose. These concerns could lead to recommendations for long-term monitoring in patients.

Mean eteplirsen plasma exposures (AUC) at the no observed adverse effect levels (NOAELs) for monkeys and juvenile rats were 20- and 6-fold, respectively, higher than that of patients who received the to-be-marketed dose of 30 mg/kg/week by the intravenous route.

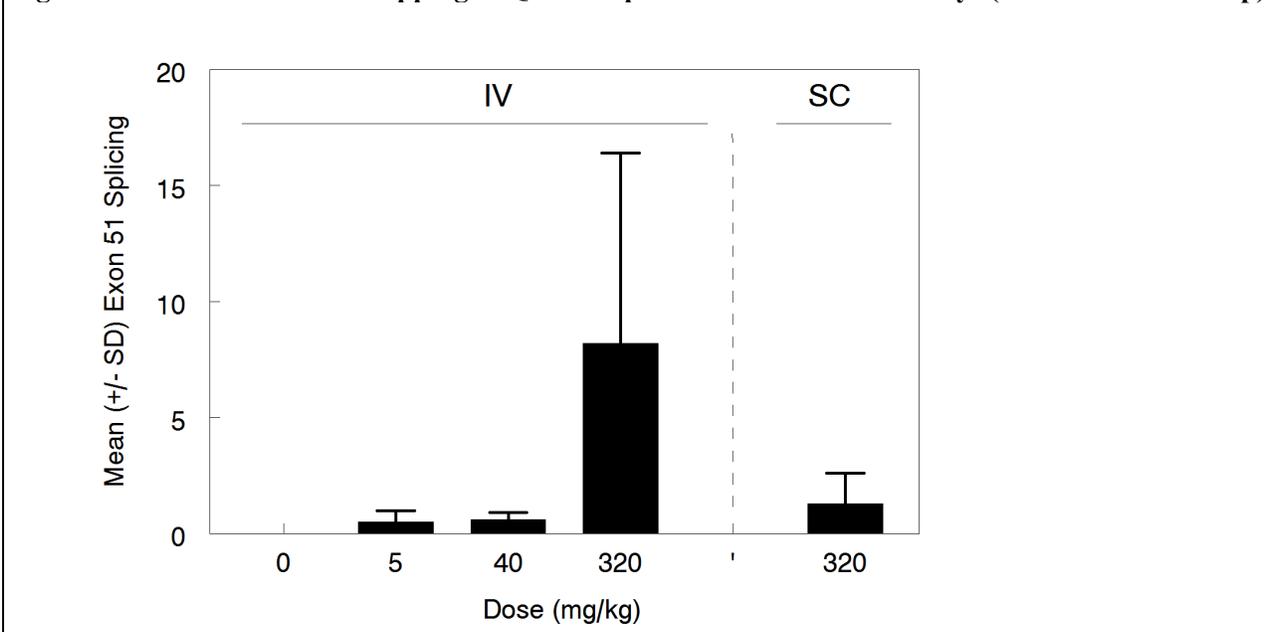
The applicant presented data on the exon skipping activity of eteplirsen in cynomolgus monkeys (“Exon skipping activity of AVI-4658 in cynomolgus monkey tissue samples from ^{(b) (4)} study ^{(b) (4)},” applicant study 4658-ssa-005). Samples of quadriceps muscle, heart, and diaphragm tissues were collected on Day 79 from cynomolgus monkeys after 12 weekly doses of eteplirsen at 0, 5, 40, or 320 mg/kg IV, or 320 mg/kg SC. Muscle samples were analyzed for exon 51 skipping of the dystrophin gene using polymerase chain reaction (PCR).

Exon skipping was detected in a nonlinear, dose-dependent manner (Table 1, Figure 1). With a 1-log increase in dose (from 5 to 40 mg/kg), there was little change in exon 51 skipping. With a second log increase in dose (from 40 to 320 mg/kg), however, there was more than a log increase in response. As noted below, the applicant studied doses of 30 and 50 mg/kg/week in the clinic (6 patients at each dose), and there is significant question as to whether the plateau of the dose-response curve was reached. It is possible that much higher doses could lead to substantially greater effects on dystrophin production – effects that could be important for efficacy.

Table 1: Average Percentage of Exon 51 Skipping in Intact Monkeys (N=8 for Each Group)

Tissue	Average % Exon 51 Splicing ± SD				
	0 mg/kg IV	5 mg/kg IV	40 mg/kg IV	320 mg/kg IV	320 mg/kg SC
Quadriceps muscle	0.0 ± 0.0	0.5 ± 0.5	0.6 ± 0.3	8.2 ± 7.4	1.3 ± 0.5
Heart	0.0 ± 0.0	0.1 ± 0.1	0.2 ± 0.2	4.5 ± 2.9	1.4 ± 0.5
Diaphragm	0.0 ± 0.0	0.2 ± 0.2	0.9 ± 0.7	6.1 ± 3.5	2.2 ± 0.9

Figure 1: Evidence of Exon Skipping in Quadriceps Muscle in Intact Monkeys (N=8 for Each Group)



With respect to the advisability of evaluating higher doses in humans, this subject is well summarized by Dr. Bastings in his Division Memo: “Considering the seriousness of DMD, the unmet medical need, and the nature of the toxicities observed in animals, I believe that the nonclinical data would support, with proper monitoring, dosing in DMD patients at least up to 200 mg/kg, a dose expected to provide exposure similar to the most sensitive species NOAEL for the toxicities seen in animals. If the human safety experience at these doses is acceptable, further dose escalation is possible in DMD patients.”

Finally, the nonclinical review team provided insight that is relevant for the interpretation of clinical data with respect to production of dystrophin protein: “The most robust finding among the studies provided and referenced in this submission was the wide variability in the extent of PMO-induced dystrophin expression within a single muscle and among different muscles, suggesting that caution is warranted in generalizing from the results of biopsies taken from only one or a few sites, muscle types, or patients.”

Carcinogenicity:

Carcinogenicity studies have not been conducted with eteplirsen. The nonclinical review team opined that carcinogenicity studies in 2 species should be conducted as a post-marketing requirement. Dr. Bastings agrees, and I agree, that for this serious indication with unmet need, carcinogenicity studies can be deferred until after marketing.

5. Clinical Pharmacology

The Clinical Pharmacology team does not recommend approval; they recommend generation of robust evidence of effectiveness prior to approval. Specifically, the team is recommending a double-blind, placebo-controlled study in patients with mutations amenable to exon-51 skipping who are likely to be ambulant for 1 year, with use of appropriate endpoints based on upper or lower body strength in patients between 4 and 12 years of age. They also suggest study of doses greater than 50 mg/kg administered weekly, or alternate regimens that would include loading and maintenance doses, for example, twice-weekly administration for 6 months followed by weekly administration for 6 months. Their recommendations are based on the 3- to 4-hour half-life of the drug, urinary excretion of 60-70% of the drug within 24 hours, and the absence of known toxicity at doses of 50 mg/kg. The immunogenicity of eteplirsen can be further assessed in future clinical trial(s) as well.

Summary of Pharmacokinetics:

- Pharmacokinetics was approximately dose-proportional and linear from 0.5 to 50 mg/kg/week, with insignificant accumulation in this dose range.
- Following single or multiple intravenous infusions, peak plasma concentrations (C_{max}) occurred near the end of infusion.
- Plasma concentration-time profiles showed multi-phasic decline, with virtually all drug eliminated within 24 hours (24 hours after completion of infusion, eteplirsen concentrations were 0.02% of C_{max}).
- At doses of 30 and 50 mg/kg, the elimination half-life is ~3.5 hours, with ~65% of the drug excreted unchanged in the urine. The drug is not metabolized.
- Protein binding of eteplirsen in humans is relatively low, ~6% to 17%, and is independent of concentration.

- The volume of distribution data suggest distribution or cellular uptake into peripheral tissues.
- Inter-subject pharmacokinetic variability is moderate, generally in the range of 20 to 55% for exposure measures (C_{max} and AUCs) as well as other key pharmacokinetic parameters.
- Intrinsic factors were not studied (typically, in a larger development program, age, gender, body weight, geographic region, hepatic impairment, renal impairment, and other potentially significant covariates would be studied).
- *In vitro* investigations on major CYP isozymes and transporters did not reveal the need for additional investigation in humans.
- Eteplirsen was not a significant inhibitor or inducer of CYP.
- Eteplirsen was not a substrate or inhibitor for any of the key human transporters tested.
- Eteplirsen is expected to have a low potential for drug-drug interactions.

Finally, the clinical pharmacology team noted that if eteplirsen were found to be safe and effective, it would likely benefit all mutations amenable to exon-51 skipping and should be labeled accordingly.

QT Effects:

QT effects were not formally investigated in man.

6. Clinical Microbiology

Not applicable.

7. Clinical/Statistical Efficacy

Sarepta is seeking accelerated approval for eteplirsen for the proposed indication:

“EXONDYS 51 is indicated for the treatment of Duchenne muscular dystrophy (DMD) in patients who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping. This indication is approved on an intermediate endpoint demonstrating delayed disease progression as measured by the 6 minute walk test [see *Clinical Studies (14)*]. Continued benefit will be evaluated through confirmatory trials.”

In this section, I provide an explanation of how accelerated approval might be used as a potential pathway to approval, based on production of dystrophin in skeletal muscle. I then discuss the evidence that eteplirsen produces dystrophin in skeletal muscle, based on immunohistochemistry and Western blot analyses. Finally, I discuss the clinical data that could serve as the basis for a conventional approval.

Accelerated Approval:

The applicant has requested accelerated approval based on an endpoint of 6-minute walk distance. The proposed indication states that 6-minute walk test is considered to be an intermediate endpoint demonstrating delayed disease progression.

There is little in the NDA to explain the applicant’s thought process here. In Sarepta’s briefing materials for the April 25, 2016 meeting of the Peripheral and Central Nervous System Drugs Advisory Committee, they stated (page 16):

“The accelerated approval pathway means that there will be an acceptable degree of uncertainty about whether the therapy will actually result in the anticipated clinical benefit. This uncertainty is addressed by the requirement that ‘appropriate post-approval studies to verify and describe the predicted effect’ would usually be underway at the time of approval.”

The applicant appears to misconstrue the intent of the accelerated approval pathway. They purport to show that, after 36 months of treatment, eteplirsen improves physical performance as assessed by the 6-minute walk test. We consider the 6-minute walk test to be a valid and meaningful measure of how well a patient functions – i.e., a clinical endpoint that would be a basis for full approval – not a surrogate endpoint or an intermediate endpoint. For slowly progressive diseases, an intermediate clinical endpoint, a clinical endpoint that can be measured earlier than an effect on irreversible morbidity or mortality and is considered reasonably likely to predict the drug’s effect on irreversible morbidity or mortality or other clinical benefit, can be used to support accelerated approval. But all would agree that showing an improvement on a clinically meaningful endpoint at 36 months would be adequate to support a conventional approval in DMD, a position we have taken with other DMD drugs.

Thus, the applicant has provided study results that purport to show improvement in a meaningful clinical endpoint after a relatively long duration of treatment, but they appear to propose accelerated approval as a means to deal with uncertainty about whether the therapy has actually been shown to provide a clinical benefit in the trial.

Clearly, if the review team had reached the conclusion that the applicant had provided substantial evidence of an effect on 6-minute walk distance during some 3 to 3.5 years of treatment, they would recommend a conventional (full) approval, and not accelerated approval. As noted in the reviews, however, for a number of reasons the review team does not believe that the applicant has provided substantial evidence of an effect on 6-minute walk distance, or any measure of physical performance (see below). Importantly, accelerated approval is not intended to enable use of less than substantial evidence of a treatment effect as a basis for approval, to be bolstered by more compelling evidence to be developed in the post-marketing setting.

Despite the lack of substantial evidence of clinical efficacy from Study 201/202 (see below), it is important to consider whether accelerated approval, based on an effect on a surrogate endpoint, could provide a viable alternative pathway to approval. 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.

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 3) whether the effect demonstrated meets the test of being “reasonably likely” to predict clinical benefit.

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 the demonstration of effectiveness, but this is not correct. 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 requirements 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).”

Again, the critical issue here is whether factor 3 (above) is met, in light of these considerations.

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. (Of note, the best-case scenario for eteplirsen is the production of an abnormal Becker-type dystrophin, not normal dystrophin, but that will be discussed later.)

The second part of factor 3 is whether an effect has been demonstrated, and the standard remains ‘substantial evidence’ based on adequate and well-controlled clinical investigations. Typically, such evidence would be two studies, both achieving a p -value < 0.05.¹

¹ 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. See: “Guidance for Industry Providing Clinical Evidence of Effectiveness for Human Drug and Biological Products;” May, 1998.

The third part of factor 3, the determination that the demonstrated effect is “reasonably likely” to predict clinical benefit, is a matter of judgment. Thus, once there is substantial evidence of a treatment effect, the determination of whether the effect size is “reasonably likely” to predict clinical benefit is an area where flexibility can be applied. Presumably there is some threshold effect that would have to be achieved in order to satisfy this criterion, but this is not described in the regulations.

Is There a Basis for Accelerated Approval: Production of Dystrophin Protein in Skeletal Muscle?

The applicant assessed skipping of the messenger RNA exon using reverse transcriptase polymerase chain reaction (RT-PCR), a standard laboratory technique to detect RNA expression. Exon 51 skipping was confirmed by RT-PCR analysis in all patients treated with eteplirsen, establishing proof of concept that eteplirsen can cause at least some degree of exon 51 skipping, as intended. Because PCR is a highly sensitive technique that can detect even a few copies of messenger RNA, the findings do not support efficacy.

Dystrophin production was assessed by two widely-used and complementary methods: immunofluorescence (immunohistochemistry) and Western blot. Immunofluorescence is generally used to assess the presence or absence of proteins in tissue sections, and is particularly useful for cellular localization of protein (by light microscopy). Western blot provides quantitative analysis of protein, but no information on cellular localization.

Originally, the applicant evaluated the effect of eteplirsen on dystrophin expression in Studies 28, 33, and 201/202.

Of note however, as the May 26, 2016 goal date was approaching, the Office of New Drugs (OND) and the Center for Drug Evaluation and Research (CDER) could not reach agreement on the regulatory action for this NDA: the Office of New Drugs favored issuance of a complete response whereas CDER favored approval.

Thus, in order to obtain definitive data on dystrophin production to support accelerated approval, we requested that the applicant perform Western blot analyses of skeletal muscle biopsy samples that had been obtained in the ongoing Study 301 (PROMOVI). The applicant was told by CDER that if they were “...successful in showing, to FDA’s satisfaction, a meaningful increase in dystrophin by Western blot analysis between the paired pre-and post-treatment samples, we expect to be able to grant an accelerated approval....” Thus, data from Study 301 were included in this NDA and discussed below.

A. Immunohistochemistry

The applicant used immunohistochemistry in cross-sections of skeletal muscle biopsies to distinguish and count “dystrophin-positive” and “dystrophin-negative” muscle fibers. The methods are described in detail in Dr. Breder’s review. Briefly, following immunostaining of tissue sections for dystrophin, 4 fields were manually selected from the 4 quadrants of each slide, and images were captured (digitized) at 20X magnification. The contrast of each image was manipulated to enhance background staining so that most of the muscle fibers became visible, making it possible for the reader to perform a manual count of the total number of fibers. Image contrast was returned to normal, and positive fibers – fibers with at least some degree of

positive staining – were manually counted. For each field, the number of positive fibers was divided by the total number of fibers to calculate the percentage of positive fibers. Various rules were prospectively established to define “positive” fibers; in essence, a fiber could be classified as “positive” if its staining intensity was only slightly perceptible over background. Importantly therefore, a reading of 50% “positive” fibers in a tissue field is not tantamount to 50% (normal) dystrophin. A 50% figure means only that half the fibers exhibited staining that was at least barely perceptible over background.

Immunofluorescence data were also analyzed using Bioquant software. For these analyses, the user determined a brightness threshold for each digitized image, in essence selecting all pixels where staining intensity exceeded a particular user-selected value. Once selected, the software calculated the mean intensity of the selected pixels. Given that the region of interest for these analyses was limited to the pixels that exceeded a threshold rather than the total image, I do not consider the Bioquant analyses to be readily interpretable.

Study 33 was a 7-patient, exploratory, phase 1 study, initiated in 2007 at the Hammersmith and Saint Mary’s Hospitals, London, UK. Two subjects received a single 0.09-mg dose of eteplirsen in the extensor digitorum brevis (EDB) muscle of one foot and placebo in the contralateral foot. Five subjects received a single 0.9-mg dose of eteplirsen in the EDB muscle of one foot and placebo in the contralateral foot. After 14 to 28 days, dystrophin was detected adjacent to the needle tracks by immunohistochemistry and Western blot. Western blot analyses were not carried out for control muscles injected with placebo.

Study 28 was a 19-patient, exploratory, phase 1 study, initiated in 2009 at 2 sites in the UK. Patients had DMD amenable to exon 51 skipping. Eteplirsen was administered weekly by the intravenous route for 12 weeks at doses ranging from 0.5 to 20 mg/kg. There were up to 4 patients per dose level. After FDA expressed concerns about the reliability of the procedures and methods, the applicant responded that “Study 28 was an exploratory phase 1b study which was only intended to generate proof of concept data to guide future studies. For this reason, quality controls for the dystrophin data in Study 28 were not properly optimized.” Some data were missing, and after considering all of this information, the review team did not deem the results to be interpretable.

Study 201 was a single-center, double-blind, placebo-controlled, parallel-dose study in 12 patients with DMD. The study was initiated in 2011. Patients were randomized (1:1:1) to eteplirsen 30 mg/kg/week, eteplirsen 50 mg/kg/week, or placebo (n=4 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. The extension trial is well described in other reviews.

The 1° endpoint of Study 201 was the percentage of dystrophin-positive fibers in muscle biopsies as assessed using immunohistochemistry. The main comparison was planned to be the 50 mg/kg/week group at Week 12 and the 30 mg/kg/week group at Week 24 to the combined placebo group. The applicant’s original results are shown in Table 2, adapted from their clinical study report. As will be noted below, these results are not deemed to be reliable.

Table 2: Adapted From Table 11-1 of Applicant’s Clinical Study Report: Effect of Eteplirsen on Dystrophin-Positive Fibers Detected by Immunohistochemistry with MANDYS106

Time point		Placebo N = 4	30 mg/kg/wk Eteplirsen N = 4	50 mg/kg/wk Eteplirsen N = 4
Baseline	Mean	15.64	18.19	11.00
	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^b	Mean	11.59	41.14	11.79
	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 ^c	0.79
	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

It should be stressed again that the figures in the table represent the percentage of dystrophin-positive fibers, but in no way correspond to the percentage or quantity of dystrophin relative to a normal individual. Muscle fibers displaying virtually any staining intensity above background were considered “positive.” As noted above, therefore, a reading of 50% positive fibers means only that 50% of fibers exhibited staining that was perceptively above background.

These results were substantially reported in a 2013 publication,² which claimed that eteplirsen markedly increased functional dystrophin production: “...*the percentage of dystrophin-positive fibers was increased to 23% of normal; no increases were detected in placebo-treated patients ($p \leq 0.002$). Even greater increases occurred at week 48 (52% and 43% in the 30 and 50 mg/kg cohorts, respectively), suggesting that dystrophin increases with longer treatment. Restoration of functional dystrophin was confirmed by detection of sarcoglycans and neuronal nitric oxide synthase at the sarcolemma.*” The publication also stated that dystrophin expression was confirmed by Western blot, with a figure showing what were termed “representative” results.

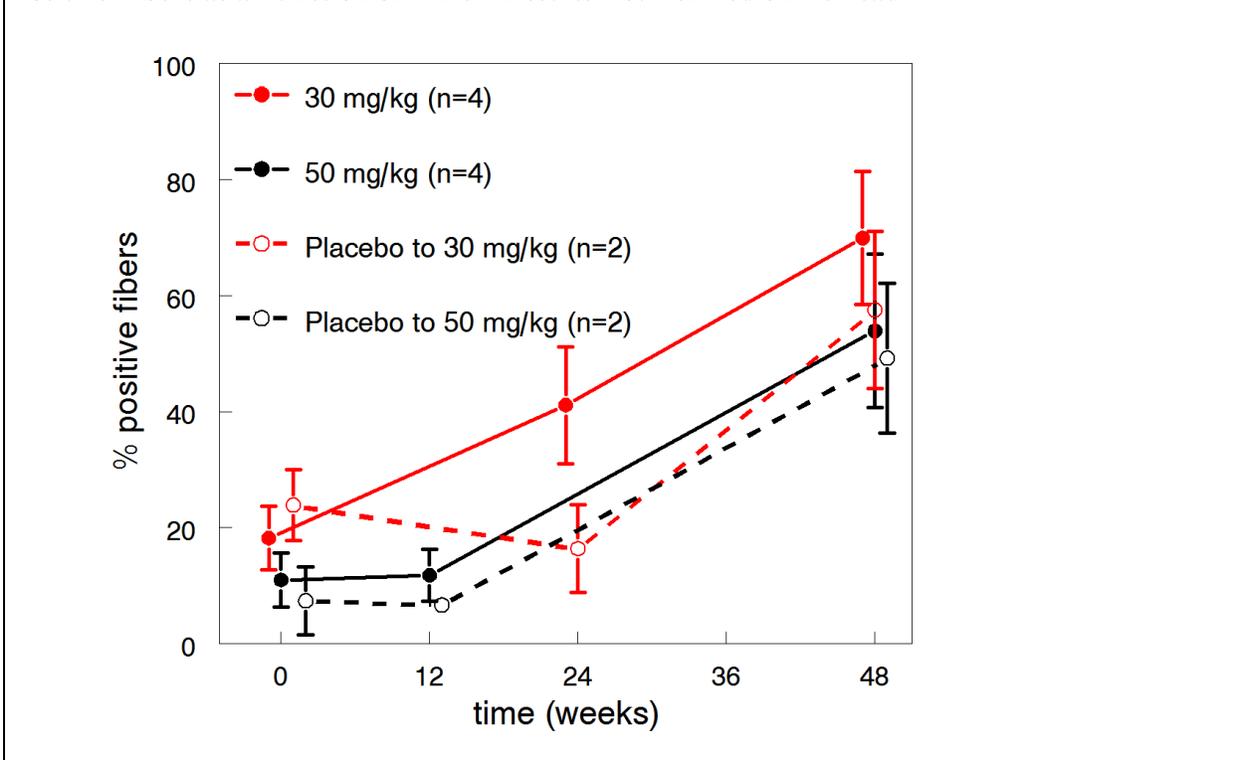
Publication of this paper was followed by a Sarepta press release,³ which also claimed a remarkable treatment effect from eteplirsen and raised wildly unrealistic expectations in the DMD community. It was these perceptions and expectations that led the applicant to declare that a placebo-controlled study was no longer feasible (see below).

² Mendell JR, et al: Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* 2013;74:637-47

³ Sarepta press release, 8/8/13, at <http://investorrelations.sarepta.com/phoenix.zhtml?c=64231&p=irol-newsArticle&ID=1846052> [downloaded 5/10/16]

The original data from Nationwide Children’s Hospital submitted to FDA are plotted in Figure 2. Immunostaining for dystrophin appears to increase markedly in all 4 groups with time, with some 50 to 60% of fibers staining positive for dystrophin at 48 weeks. For reasons explained below, the review team disagrees with the veracity of these data.

Figure 2: Original Results of Dystrophin Immunostaining Using MANDYS106 Antibody: Percent Positive Fibers as a Function of Time – Results Not Verified on Re-read



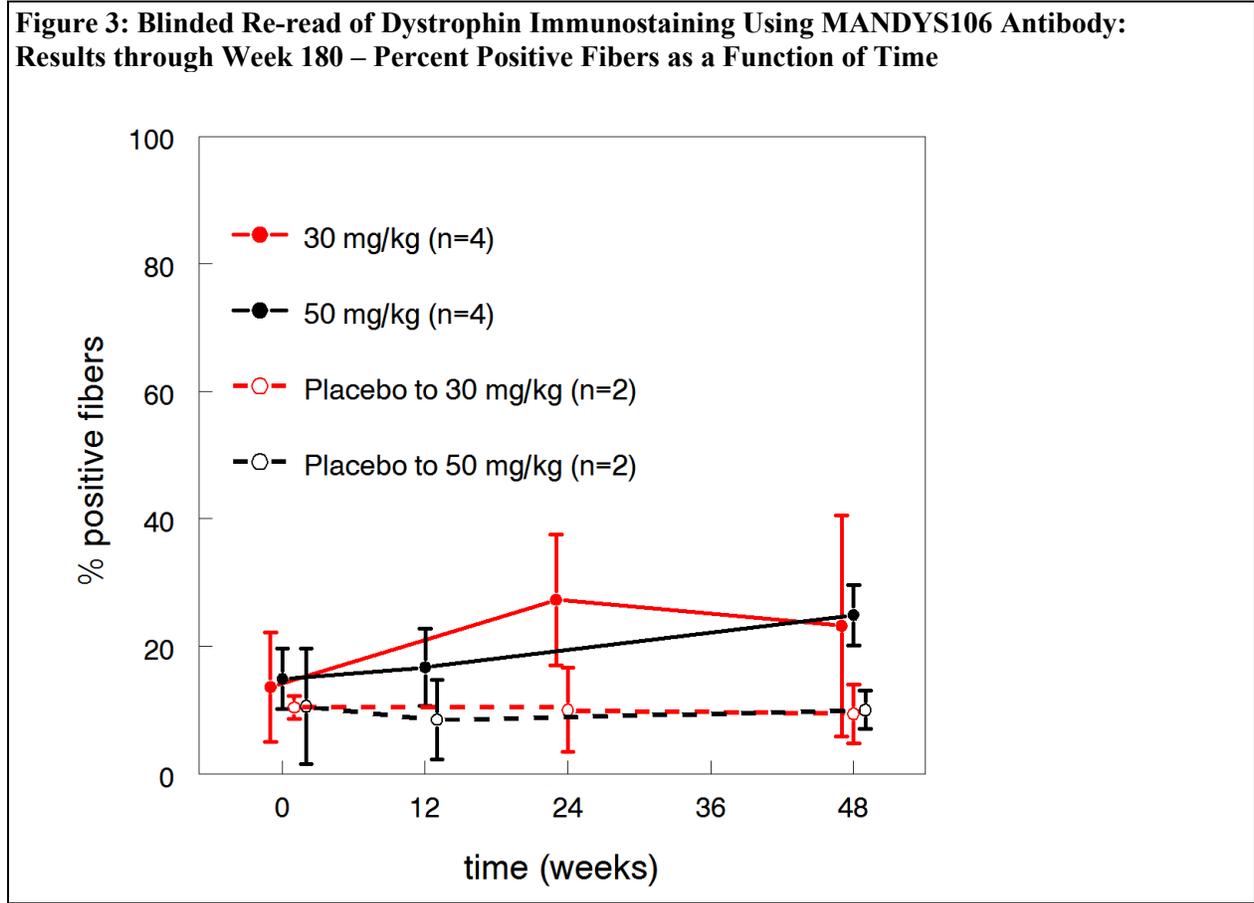
I was part of an inspection team that conducted (May 29 and 30, 2014) a site visit to Nationwide Children’s Hospital in Columbus, OH, where Study 201 was conducted. We found the analytical procedures to be typical of an academic research center, seemingly appropriate for what was simply an exploratory phase 1/2 study, but not suitable for an adequate and well controlled study aimed to serve as the basis for a regulatory action. The procedures and controls that one would expect to see in support of a phase 3 registrational trial were not in evidence.

Although the technician had been blinded to treatment group, access to the treatment code was not protected with the kinds of safeguards and firewalls that one would ordinarily put in place for an adequate and well controlled trial. The immunohistochemistry images were only faintly stained, and had been read by a single technician using an older liquid crystal display (LCD) computer monitor in a windowed room where lighting was not controlled. (The technician had to suspend reading around mid-day, when brighter light began to fill the room and reading became impossible.) These issues are well described in a summary of inspectional findings in Dr. Breder’s clinical review (page 27). There was also concern that the reader, although masked to treatment assignment, was not masked to sequence/time (see below). Importantly, in a trial where all patients eventually received the active drug, knowledge of sequence could lead to the false appearance of a treatment effect, i.e., the appearance of increasing dystrophin expression

with time, simply by having a lower threshold for calling fibers “positive” at later time points in the study.

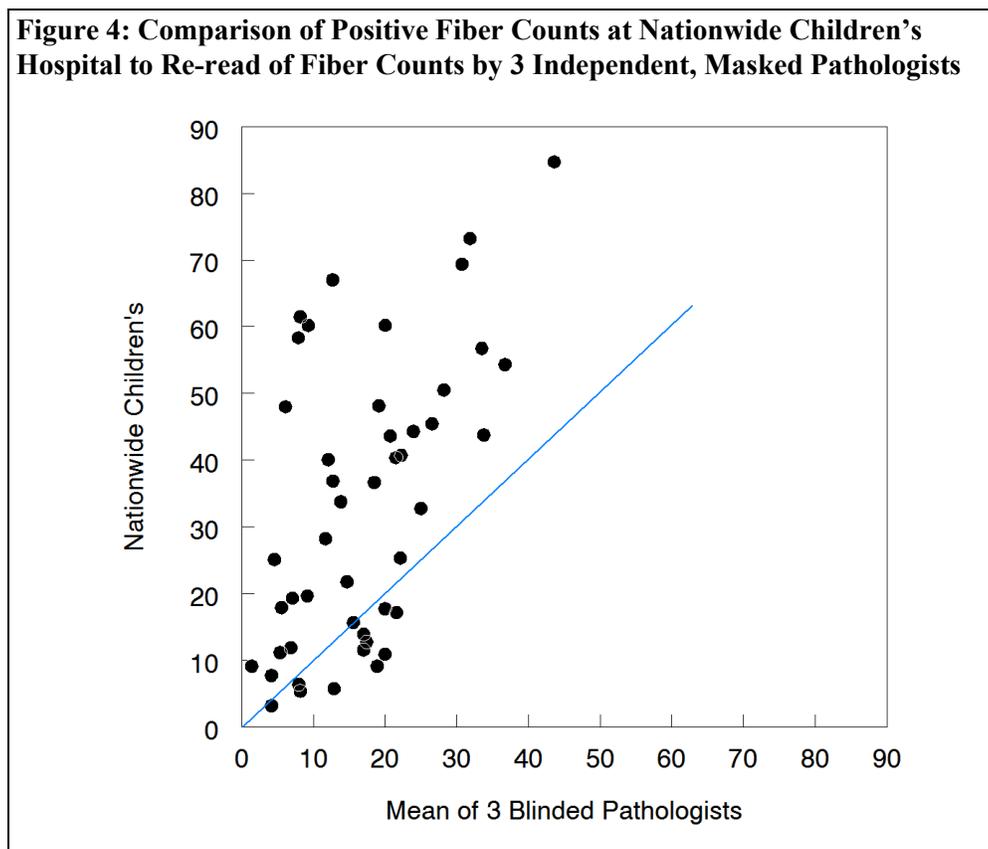
Having uncovered numerous technical and operational shortcomings in Columbus, our team worked collaboratively with the applicant to develop improved methods for a reassessment of the stored images. We suggested a re-read of all images by 3 independent masked readers, such that blinding could be assured and inter- and intra-observer variability could be characterized. We also suggested the use of better equipment, specifically, high-quality light-emitting diode (LED) computer monitors, in darkened rooms.

The applicant undertook a blinded re-analysis of the images on the server as FDA suggested. Unfortunately, the re-analyses failed to show a significant increase in dystrophin-positive fiber counts in eteplirsen treated patients (Figure 3). Note also that for patients who switched from placebo to eteplirsen at Week 24 (dashed red and black lines), there was no response between Weeks 24 and 48.



This re-analysis, along with the study published in 2013,² provides an instructive example of an investigation with extraordinary results that could not be verified. The publication, now known to be misleading, should probably be retracted by its authors.

Figure 4 shows the correlation between the dystrophin immunohistochemistry data as read by the technician at Nationwide Children’s Hospital and the 3 blinded pathologists. Each point represents data from a single patient at a single time point (an analysis of 24 images), as read by Nationwide Children’s Hospital (y-axis) and the group of 3 blinded pathologists (x-axis). Readings from the 3 pathologists are averaged. Perfectly correlated readings would lie along the blue line of unity. In most cases, the reading from Nationwide exceeds the reading from the pathologists, i.e., above and to the left. Thus, despite less-than-optimal lighting conditions that should have favored reduced reading of positive fiber counts at Nationwide Children’s Hospital, there was a striking tendency for the reporting of higher counts at that institution.

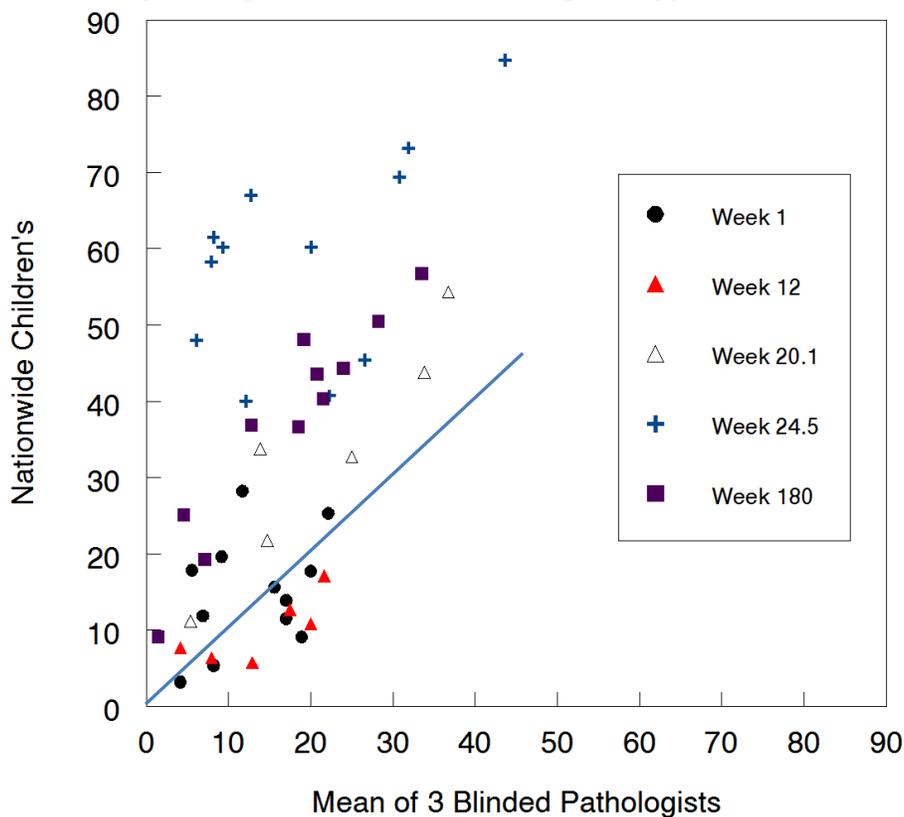


One might reasonably ask why the original readings were not reproduced by a blinded re-read. Figure 5 shows the same scatterplot between readings by Nationwide Children’s Hospital and the 3 blinded pathologists. In this display, however, readings from samples obtained at the disparate time points are shown with unique markers.

It is striking that the deviations between the readings of Nationwide and the re-read by the blinded pathologists differ substantially by study time point. Thus, at Week 1 (●) and Week 12 (▲), time points before increased dystrophin production would be expected, there is reasonable agreement between Nationwide and the pathologists, i.e., the points lie close to the blue line. In contrast, for the Week 24.5 time point (+), readings from Nationwide Children’s Hospital are much higher than those of the 3 pathologists, suggesting that blinding to sequence (i.e., time

point) was not achieved. At the time the Week 180 samples were read at Nationwide Children's Hospital, the technician was aware that the images would be re-read by 3 pathologists, which could explain why there is less exaggeration (i.e., the Week 180 readings are closer to the blue line of unity than the Week 24.5 readings).

Figure 5: Comparison of Positive Fiber Counts at Nationwide Children's Hospital to Re-read of Fiber Counts by 3 Independent, Masked Pathologists: Apparent Interaction with Time



Week 180 Data

As noted by the review team, the extension phase of the study (Study 202) has continued through the present. Eleven (11) of the 12 patients consented to undergo a fourth skeletal muscle biopsy at Week 180 (3.5 years), and these samples were analyzed using immunohistochemistry.

Prior to the analysis of the Week 180 samples, however, the applicant made changes to the immunohistochemistry protocol with the intent of decreasing non-specific staining. Their aim was to compare the Week 180 dystrophin level to baseline for each patient. Frozen archived baseline tissue was available for only 3 of the patients, however, and so the applicant supplemented these with samples from 6 untreated external DMD patients, all to be compared to the Study 201/202 patients at Week 180. Images were read by the same 3 pathologists, masked to treatment group.

For this analyses, the applicant claims a remarkable increase in dystrophin staining: the 9 baseline samples (including samples from 3 patients in Study 201/202 and 6 external controls) showed a mean percent positive fiber count of $1.1 \pm 1.3\%$ (mean \pm SD), whereas the Week 180 samples showed a mean percent positive fiber count of $17.4 \pm 10.0\%$. I will note that FDA made no attempt to inspect or oversee the new immunohistochemistry methods.

Given that the original baseline percent positive fiber count for patients from Study 201/202 was $13.0 \pm 6.2\%$, it would be important to understand why the results from a new immunohistochemistry protocol provided results more than an order of magnitude lower ($1.1 \pm 1.3\%$).

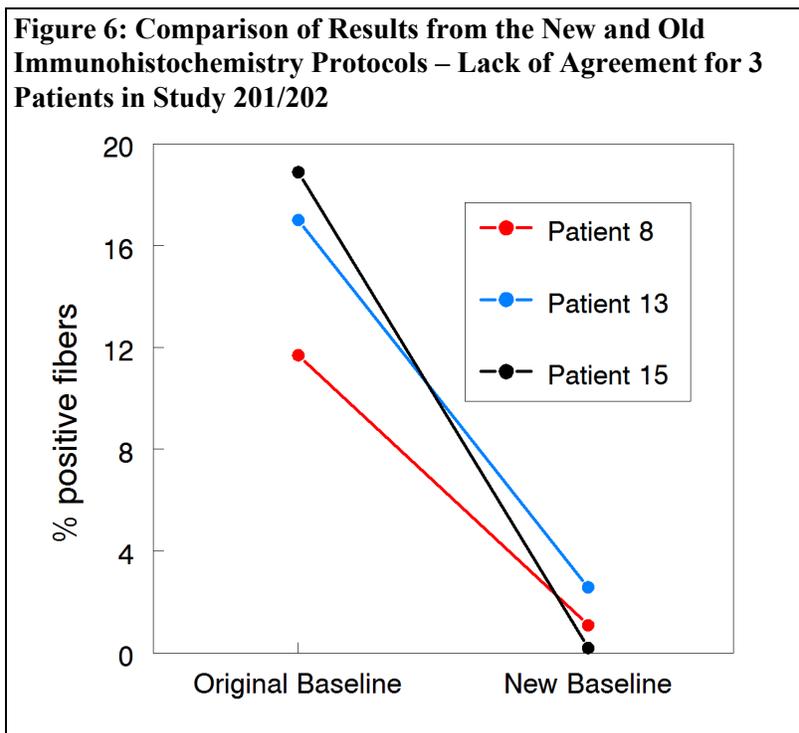
As noted above, there were 3 patients in Study 201/202 with adequate archived tissue for separate immunohistochemistry analyses using both the old and new methods. Figure 6 shows how the two methods compare. These are essentially replicate analyses of a single tissue sample using the two methods.

There is an inexplicable difference of more than an order of magnitude between results using the new and old immunohistochemistry protocols. These 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.

Commentary:

The review team provided much thoughtful discussion regarding the relative merit of immunohistochemistry for the quantitative assessment of dystrophin in skeletal muscle. My view is that such analyses, if properly blinded and controlled, can yield semi-quantitative information that could show differences in dystrophin production, e.g., 50% is more than 25%, although the method does not allow correlation of particular values of "percent positive" fiber counts with quantitative measures of muscle protein. Moreover, comparisons of fiber counts across centers, across experiments, or, for that matter, across staining or reading runs within a single laboratory, do not seem likely to be informative.



Recognizing that Study 201/202 was a small exploratory phase 1/2 study that was not powered to show a small change in dystrophin, the study provides no evidence of increased dystrophin production by immunohistochemistry.

It is unfortunate that the original readings from Nationwide Children's Hospital, purporting to show a marked effect of eteplirsen on dystrophin-positive fiber counts – counts now known to be unreliable – led to the perception that the drug produces large amounts of dystrophin. These results fueled the public perception that eteplirsen is highly effective as well as the DMD community's reluctance to participate in placebo-controlled trials. Only recently, an unauthored report in the Wall Street Journal stated: "The trial turned up evidence that eteplirsen makes good on pumping out dystrophin, a feat no treatment has managed."⁴ Presumably this misperception has been carried over from the initial 2013 reports.

B. Western blot

1) Data analyzed prior to the PDUFA goal date

A second, more important line of evidence regarding dystrophin production is Western blot, a standard, widely-used, analytical technique to assess levels of protein in biological tissues. Western blot was used to quantify dystrophin protein directly, and the methods are described by others.

For a variety of reasons discussed by Dr. Rao, the Western blot analyses originally conducted by the applicant were technically unsatisfactory. The Western blots from the first 3 time points had oversaturated bands, lacked appropriate controls, and were essentially uninterpretable. After conducting a site visit to the Columbus OH laboratory, FDA rendered advice to the applicant with the goal of improving technical aspects of the assay for future use.

The applicant amended the study protocol to allow for an additional skeletal muscle biopsy at Week 180 (3.5 years), potentially enabling pre- to post-treatment comparisons of Becker-type dystrophin after prolonged eteplirsen treatment. As noted above, 11 of the 12 patients in Study 201/202 consented to undergo a fourth skeletal muscle biopsy at Week 180. Of note, the baseline samples had been obtained from biceps muscle, whereas the Week 180 samples were obtained from deltoid muscle.

Two blocks were prepared from each patient sample. Sections from both blocks were pooled during homogenization for lysate preparation, and Western blots were run in duplicate.

The individual (anonymized) values for the Western blot analysis are shown in Table 3. As reported by the review team, the analysis for 11 of the 12 original patients showed a mean dystrophin value of $0.93\% \pm 0.84\%$ of normal (mean \pm standard deviation) after 3 to 3.5 years of eteplirsen treatment (3 years in patients initially randomized to placebo; 3.5 years in the other patients). Mean values were virtually the same for the lower (30 mg/kg/week) and higher (50 mg/kg/week) dose groups; there is no suggestion of a dose-response.

⁴ A Legal Test for the FDA: Black letter law dictates approval for a muscular dystrophy drug; Wall Street Journal, May 9, 2016.

Of note, the Western blot values are quite variable, both between patients and between duplicate runs *within* patients (i.e., repeatability; intra-assay precision), Table 3.

Mean values ranged from a maximum of 2.47% in Patient J, to near-zero in Patient H, and to zero in 2 patients (E and G). For some patients, there were considerable discrepancies between duplicate runs (the intra-assay difference was >0.5% in Patients B, C, D, and J). Aside from patients with zero or near-zero dystrophin, only 3 patients showed reasonable intra-assay agreement: Patients F, L, and K.

Given that these numbers represent duplicate runs from tissue homogenates, intra-assay differences suggest limited precision/reproducibility of the method, heterogeneity of the samples, or both.

Subject	Dose	Western blot				Group Mean ± SD
		gel 1	gel 2	Mean (arithmetic)	Mean (per protocol)	
L	30 mg/kg	0.58	0.46	0.52	0.52	
K	30 mg/kg	1.45	1.78	1.62	1.62	
J	30 mg/kg	2.83	2.11	2.47	2.47	
H	30 mg/kg	0.02*	0.28	0.15	0.14	
G	Placebo to 30 mg/kg	0.17*	0.15*	0.16	0	
F	Placebo to 30 mg/kg	0.93	1.02	0.98	0.98	0.96 ± 0.95
E	50 mg/kg	0.19*	0.16*	0.18	0	
D	50 mg/kg	0.75	0.24*	0.50	0.38	
C	50 mg/kg	1.22	0.69	0.96	0.96	
B	50 mg/kg	2.43	1.67	2.05	2.05	
A	Placebo to 50 mg/kg		1.15	1.15	1.15	0.91 ± 0.79

* below limit of quantitation

Change in Dystrophin with Treatment:

The critical question, of course, is whether the value of 0.93% is meaningfully greater than the value at baseline, or even meaningfully greater than zero. Assuming that one considers this value greater than zero, the baseline pre-treatment levels of dystrophin in these 11 patients are critical in determining whether eteplirsen was responsible for the dystrophin detected at Week 180.

Unfortunately, adequate pre-treatment tissue samples were available for only 3 of these 11 patients. Thus, the applicant supplemented these data with muscle biopsies from 6 untreated patients with DMD amenable to exon 51 skipping who were external to the study.

Whereas the Week 180 samples were obtained from deltoid muscle, 8 of 9 of the controls were obtained from biceps muscle (the other one was obtained from deltoid). As noted above, the non-clinical review team found "...wide variability in the extent of PMO-induced dystrophin expression within a single muscle and among different muscles, suggesting that caution is

warranted in generalizing from the results of biopsies taken from only one or a few sites, muscle types, or patients.” Use of disparate muscle groups between patients in Study 201/202 and controls was, obviously, ill advised. The finding of a difference between patients in Study 201/202 and the external controls could simply represent a difference between muscles.

FDA’s advice to the applicant (March 30, 2015) is still germane: "The control biopsy tissue that you propose to use is from a number of different muscle groups, such that differences that may exist in dystrophin expression among muscle groups may affect your results. However, in the context of other major sources of variability among biopsies (including both intra- and inter-individual differences even within the same muscle group), it appears reasonable for you to proceed with these controls, with the understanding that dystrophin changes would need to be robust to be interpretable as a drug effect."

Averaging Western blot data from pre-treatment biopsies of the 2 patients from Study 201/201 and the external treatment-naïve patients, the applicant reported a baseline dystrophin value of 0.08% ± 0.13% (mean ± standard deviation). Obviously, all but 2 of these controls are external, such that the comparison to the treated patients in Study 201/202 is non-randomized and indirect.

Table 4: Individual Untreated DMD Control Samples, Western Blot Analysis (% Normal Dystrophin)

Study; Subject	Dose	Western blot				Group Mean ± SD	All Mean ± SD
		gel 1	gel 2	Mean (arithmetic)	Mean (per protocol)		
201/202; X	0	0.05*	0.07*	0.06*	0	0 ± 0	0.08 ± 0.13
201/202; A	0	0.19*	0.08*	0.14*	0		
201/202; B	0	0.13*	0.07*	0.10*	0		
external; A	0	0.12*	0.14*	0.13*	0	0.12 ± 0.15	
external; B	0	0.03*	0.12*	0.08*	0		
external; C	0	0.37	failed	0.37	0.37		
external; D	0	0.04*	0.30	0.17*	0.15		
external; E	0	0.20*	failed	0.20*	0		
external; F	0	0.40	0.09*	0.25*	0.20		

* below limit of quantitation

In determining whether there is substantial evidence that eteplirsen produced dystrophin in the patients in Study 201/202, the critical questions are whether these values, near the lower limit of quantification of the assay, are actually interpretable, and whether the comparison between these subjects and a predominantly external group of untreated patients is valid.

The review team has pointed out important limitations with respect to comparability of the Western blot results from the untreated controls, summarized below:

- Biopsies from controls were obtained from biceps, whereas Week 180 biopsies from eteplirsen-treated patients were obtained from deltoid. There is some evidence that dystrophin concentrations differ by muscle group, and the study does not account for this possibility. Because the study is not well controlled, the difference between these groups of patients cannot be attributed to a drug effect.
- Two-thirds (6 of 9) of the control patients were from Study 301, and were external to study 201/202. There is no way to know how these particular patients were selected for the purpose of this comparison.
- Degradation of dystrophin or loss of immunoreactivity might occur during prolonged storage of tissue samples. If so, it could have affected the baseline samples from the 3 patients in Study 201/202, which were frozen for over 3 years prior to analysis. Note that the data are consistent with loss in immunoreactivity over time (Table 4). The per-protocol values for all 3 patients from Study 201/202 whose samples were stored for 3 years are 0 (top), whereas 3 of 6 of the samples from the external controls (bottom) are greater than zero. Although the numbers of samples are small and the comparison is non-randomized, the data nevertheless support the concept that immunoreactive dystrophin decreases during storage.

For these reasons, the review team questioned the comparability of these two groups of patients, and I agree. Having compared samples from different muscle groups in independent groups of patients, the study was not adequate and well controlled; therefore, the validity of the comparison is uncertain. The 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).

Having heard arguments and opinions from both the applicant and the review team, the Advisory Committee, despite extraordinary public activism and pressure to vote favorably, voted 7 to 6 that the applicant had not provided substantial evidence from adequate and well controlled studies that eteplirsen induces production of dystrophin to a level that is reasonably likely to predict clinical benefit. Moreover, 2 of the Committee members who voted “yes” were patient representatives.

Correlation between the applicant’s two methods to assess dystrophin

The discussion of the Week 180 dystrophin analyses would not be complete without a comparison of the results of the two complementary methods used by the applicant. Of note, the improved immunohistochemistry analyses and Western blot analyses were performed on *the same blocks of tissue*, and one should expect a reasonable correlation between the two methods if in fact the data are reliable.

Of note, there is a striking lack of correlation between these two methods of dystrophin assessment (Figure 7). It is simply not possible to determine whether the immunohistochemistry methods are inaccurate, whether the Western blot methods are inaccurate, or whether both methods are inaccurate. My view is that it is not possible to render a positive regulatory decision on the basis of unreliable data from these 11 patients. Internal consistency is lacking.

2) Data analyzed after the PDUFA goal date

As noted above, as the May 26, 2016 goal date was 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. As described by Drs. Rao, Farkas, and Bastings, Western blot analyses were performed on paired biceps samples from 13 of the patients. For each of these patients, samples had been obtained at baseline (prior to treatment) and at Week 48, after 48 weekly infusions of eteplirsen 30 mg/kg.

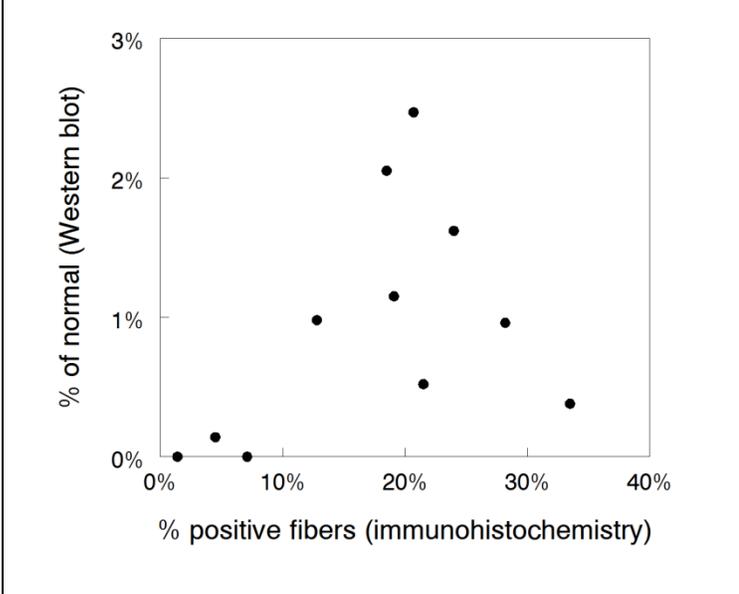
The age of these 13 patients ranged from 7 to 13 years. Paired pre- and post-treatment samples were run in side-by-side lanes on the gels, and each gel was run in duplicate. A muscle sample from a healthy 14 year-old boy with no pathologic diagnosis served as the reference sample; values from the DMD patients were reported as percent of normal.

Dr. Ashutosh Rao from the Office of Biotechnology Products reviewed the methodology and the technical reliability of the Western blot assay. Dr. Rao also conducted an inspection with Young Moon Choi, Ph.D. (Office of Study Integrity and Surveillance) and Mark Babbit (Office of Regulatory Affairs) as the analyses were being run. Xiang Ling, Ph.D., from the Office of Biostatistics, performed the statistical review on the data.

According to the protocol, acceptance of the result from each gel was contingent on two factors: 1) the R^2 value for the linearity of the standard curve of the normal control had to be > 0.9 ; and 2) the dystrophin band for the negative control DMD sample on the gel had to have a density lower than the lowest sample of the standard curve (0.25%). Samples that did not meet both criteria were deemed 'failed' and were not considered in the analyses. As it turned out, 22 of the 52 gels (42%) failed, such that many of the values represent single readings rather than the average of two. There was one patient for whom none of the values met acceptance criteria. Thus, the applicant reported pre- and post-treatment data for 12 of the 13 patients.

The applicant used 3 methods to consider values below the 0.25% lower limit of quantification: 1) consider such values to be zero; 2) analyze such values as actually reported; and 3) consider such values to be 0.24%.

Figure 7: Study 201/202 Week 180 Dystrophin Assessment – Lack of Agreement between Immunohistochemistry and Western Blot



The review team believes the most appropriate analysis is the second: analysis of all values as reported, but the results were similar for all 3 methods.

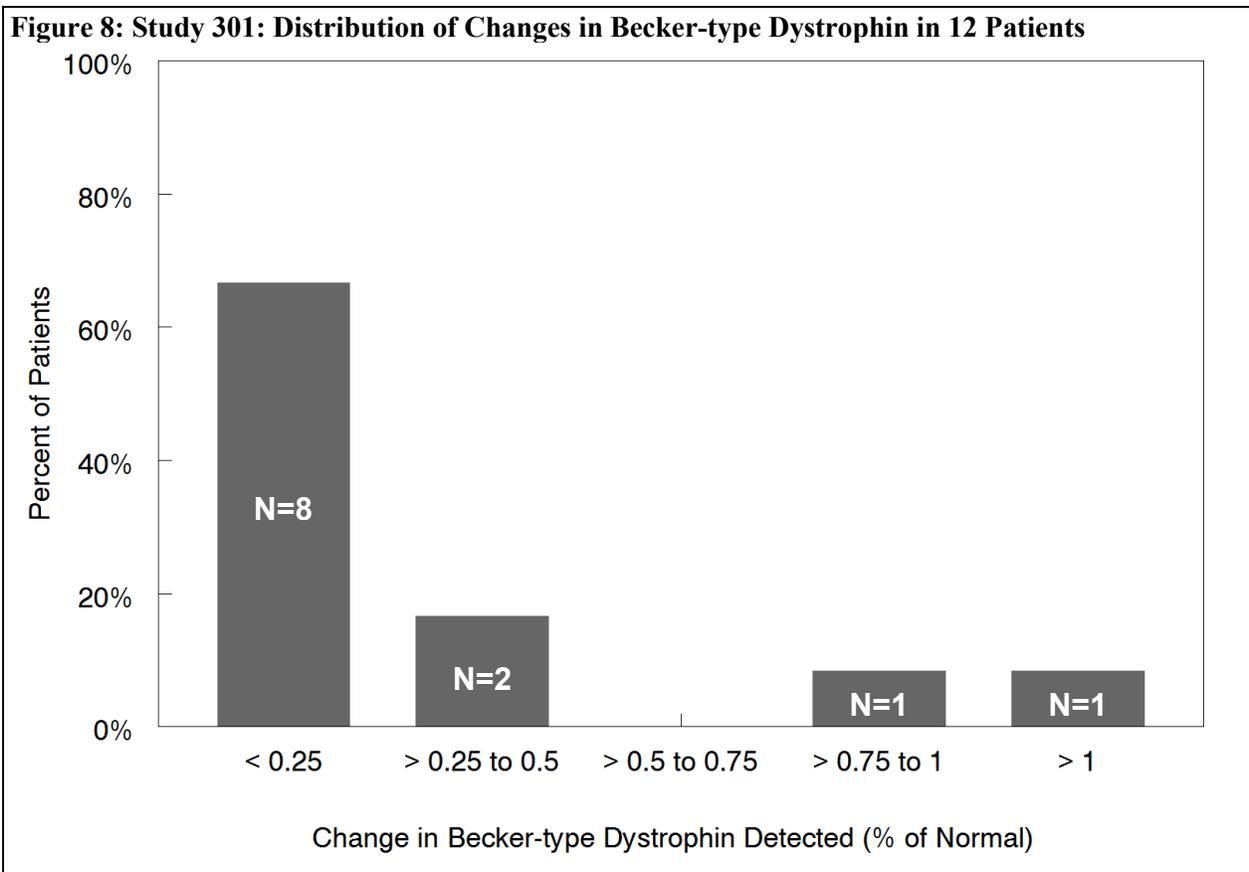
Reporting values below the limit of quantification as 0, pre- and post-treatment values are 0.06% ± 0.14% and 0.38% ± 0.50%, respectively (mean ± standard deviation), $p < 0.05$. For the 'as reported' analysis, pre- and post-treatment values are 0.16% ± 0.12% and 0.44% ± 0.43%, respectively, $p < 0.05$. Reporting all values below the limit of quantification as 0.24%, pre- and post-treatment values are 0.26% ± 0.05% and 0.48% ± 0.41%, respectively, $p < 0.05$. Individual data for the 'as reported' analysis are shown in Table 5, adapted from listing 1.1 of the applicant's "Preliminary Report: Western Blot Interim Analysis of Novel Dystrophin Expression in Muscle Biopsy Samples from Week 48 of the Clinical Study 4658-301," submitted June 27, 2016.

Irrespective of the method used to express data below the limit of quantification, the mean change is similar, ranging from 0.22% to 0.32% of normal, a treatment effect of approximately 2 to 3 parts per thousand.

Table 5: Study 301: Pre- and Post-treatment Values of Becker-Type Dystrophin

Patient	Time	status	value (%)	mean (%)	delta (%)	Patient	Time	status	value (%)	mean (%)	delta (%)
1	Baseline	pass	0.15	0.13	0.13	8	Baseline	fail	0.08	0.24	1.33
		pass	0.11					fail	0.14		
	Week 48	pass	0.22	0.26			Week 48	fail	0.08	1.57	
		pass	0.29					fail	0.05		
2	Baseline	pass	0.35	0.35	0.01	9	Baseline	fail	0.14	0.11	0.01
		fail	0.26					pass	0.24		
	Week 48	pass	0.36	0.36			Week 48	fail	1.17	0.12	
		fail	0.12					pass	1.57		
3	Baseline	pass	0.06	0.06	0.31	10	Baseline	pass	0.11	0.05	0.43
		pass	0.06					fail	0.05		
	Week 48	pass	0.5	0.37			Week 48	pass	0.12	0.47	
		pass	0.24					fail	0.11		
4	Baseline	pass	0.04	0.04	0.06	11	Baseline	pass	0.01	0.02	0.07
		fail	0.06					pass	0.08		
	Week 48	pass	0.1	0.1			Week 48	pass	0.31	0.09	
		fail	0.19					fail	0.01		
5	Baseline	fail	0.1	0.17	0.85	12	Baseline	pass	0.02	0.02	0.03
		pass	0.17					fail	0		
	Week 48	fail	0.92	1.02			Week 48	pass	0.09	0.21	
		pass	1.02					fail	0.21		
6	Baseline	pass	0.37	0.37	-0.07	13	Baseline	fail	0.34	0.18	0.03
		fail	0.46					pass	0.18		
	Week 48	pass	0.3	0.3			Week 48	fail	0.34	0.21	
		fail	0.29					pass	0.21		
7	Baseline	fail	0.04	0.17	0.25	13	Baseline	fail	0.34	0.18	0.03
		pass	0.17					pass	0.21		
	Week 48	fail	0.22	0.42			Week 48	fail	0.34	0.21	
		pass	0.42					pass	0.21		

The distribution of these changes is shown graphically in Figure 8. Of these 12 patients, 8 (two-thirds) had a change of 0.25% or less; only 1 patient (8%) had a treatment effect greater than 1%.



Commentary: Study 301 was a baseline-controlled study, where each patient served as his own control: pre- and post-treatment biopsies were obtained from the same muscle and Western blot analyses were run concurrently. An FDA inspection team observed the performance of the assays and considers the results to be reliable. Thus, unlike the data obtained from Study 201/202, the Study 301 data are considered by the review team to have been generated from an adequate and well controlled study. Study 301 provides substantial evidence of an effect of the surrogate endpoint – Becker-type dystrophin.

The critical question is whether the quantity of dystrophin produced here – a mean of 2 to 3 parts per thousand – is reasonably likely to predict clinical benefit.

With levels of Becker-type dystrophin higher in Study 201/202 (at Week 180) than in Study 301 (at Week 48), the applicant speculates that there is greater dystrophin accumulation with longer durations of treatment. These differences, however, could also be due to cross-laboratory methodological differences or play of chance; therefore, such an interpretation is highly speculative.

The Question of “Reasonably Likely to Predict Clinical Benefit”

As discussed above, the accelerated approval of eteplirsen hinges on: 1) whether Becker-type dystrophin is an appropriate surrogate endpoint for the disease; 2) whether there is substantial evidence that eteplirsen produces Becker-type dystrophin in skeletal muscle, and 3) whether such dystrophin produced is reasonably likely to predict clinical benefit, i.e., whether it is functional, and whether the quantity produced is adequate.

1. Is dystrophin an appropriate surrogate endpoint for Duchenne muscular dystrophy?

The review team believes that dystrophin is on the causal pathway of the disease, and there is no debate about the appropriateness of dystrophin as a surrogate endpoint for Duchenne muscular dystrophy.

2. Is there substantial evidence that eteplirsen produces dystrophin in skeletal muscle?

Prior to receiving the new Western blot data from Study 301 on June 27, 2016, the review team did not believe that substantial evidence from adequate and well controlled trials had been submitted to support an accelerated approval.

Study 201/202: Immunohistochemistry analyses were performed to assess and compare percent dystrophin-positive fibers at various time points before and during treatment. This is a standard technique that has been used by many laboratories for decades to assess dystrophin levels in DMD and Becker’s patients. Importantly, the analysis showed no evidence of dystrophin production through 48 weeks of treatment with eteplirsen. This information is particularly germane, because, unlike the Western blot analyses from Study 201/202, the immunohistochemistry analyses are adequately controlled. The lack of a positive finding from the blinded re-read of the immunohistochemistry data with proper controls undercuts the evidence of dystrophin production from Western blot analyses.

The applicant supplemented these data with new analyses from Week 180 that purport to show a remarkable increase in dystrophin from pre-treatment levels. Unfortunately, an altered immunostaining protocol was used, and there was an inexplicable difference of more than a log between results from the new and old protocols, rendering interpretation impossible.

The Western blot data from Study 201/202 were largely externally controlled, and there were questions with respect to the proper selection of control patients, differences in the specific muscles analyzed, and concerns regarding the possible degradation of immunoreactive dystrophin in tissue samples that might occur during long-term storage and lead to a false-positive result. Importantly, ignoring the baseline data and focusing only on the Week 180 samples, there is a striking lack of correlation between the immunohistochemistry data and the Western blot data, i.e., there is no internal consistency. Thus, these data provide no basis to believe that the study was adequate and well controlled.

Study 301: The new data submitted on June 27, 2016 were obtained from an adequate and well controlled study. This baseline-controlled study shows a statistically significant increase in Becker-type dystrophin with treatment, the surrogate endpoint. Thus, there are now data showing Becker-type dystrophin production, albeit at a small level, from one adequate and well controlled trial (Study 301), with inconclusive data from Study 201/202.

The question of “reasonably likely” is, therefore, an issue of the quantity of protein produced. As noted above, Study 301 showed a treatment effect of 2 to 3 parts per thousand in Becker-type dystrophin after 48 weeks. Study 201/202, although not adequate and well controlled, nevertheless suggested a treatment effect of 8 to 9 parts per thousand after 3.5 years.

3. Is the dystrophin that was produced reasonably likely to predict clinical benefit, i.e., is it functional, and is the quantity 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. Function is therefore a matter of judgment for which regulatory flexibility can be extended. The review team has been willing to assume that whatever Becker-type dystrophin is produced would function as well as in the Becker form of the disease. Although there can be no certainty on this point, the uncertainty is small relative to the uncertainty regarding the adequacy of the quantity, 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.*

It must be stated that the minimum level of Becker-type dystrophin that is reasonably likely to predict clinical benefit in patients with DMD is unknown. The raw data are shown in Figure 9, but this is an area where we must consider what is known about the disease and apply medical judgment.

There are two ways to consider the quantity of Becker-type 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.

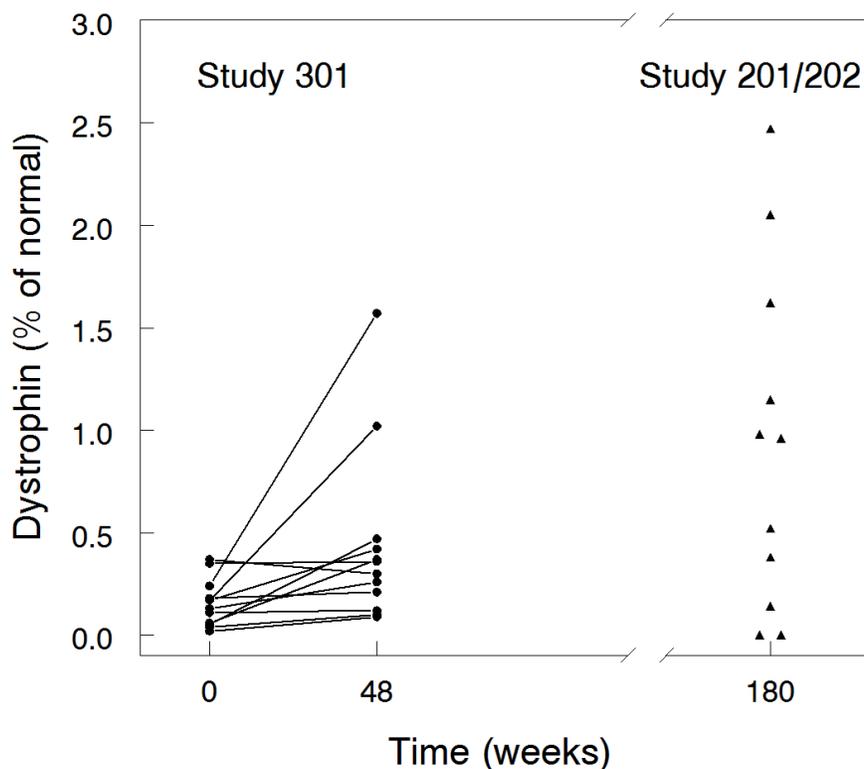
The problem with a responder analysis is that there is no rational basis upon which to define a threshold for a ‘response.’ Various cut-points could be selected, but their selection would be arbitrary, and the particular threshold chosen would have a major influence on the effect size.

Drs. Farkas and Bastings have tried to provide a framework to help put these small increases into perspective. The applicant’s data show that dystrophin levels in treatment-naïve DMD patients range from 0 to approximately 0.4% by Western blot; the applicant has not detected values > 0.4% in treatment-naïve patients.

DMD experts, including those involved with the development of eteplirsen, have stated that levels < 3% are generally associated with the typical DMD phenotype, and no patient has been found to have or produce a level of Becker-type dystrophin > 3% in response to treatment.

In order to place these small quantities of Becker-type dystrophin into a clinical perspective, many have focused on publications from a number of laboratories that attempt to relate particular levels of dystrophin protein to clinical course, e.g., maintenance of physical function, age at loss of ambulation. Some have also cited non-clinical data to relate dystrophin levels to maintenance of physical function. It is important to recognize, however, that many methodological factors affect the results of these assays, and comparison of values across various laboratories could lead to erroneous conclusions.

Figure 9: Studies 301 and 201/202: Expression of Becker-type Dystrophin by Western Blot



Van den Bergen *et al* studied the relation between dystrophin levels (quantified by Western blot) and clinical severity in 33 patients with Becker muscular dystrophy (van den Bergen JC, et al. *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 low 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.”

Chamberlain, who stated at the open public session at the advisory committee meeting that very low levels of dystrophin may be beneficial, discussed in a published paper (*Basic Appl Myol.* 7 [3&4]: 251, 1997) that “...a majority of fibers must accumulate approximately 20% of wild-type levels of dystrophin for a significant correction of the muscle pathology,” a view that seemingly contradicts the comments he made at the advisory committee meeting.

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%. Variability was particularly pronounced with low levels of dystrophin.

During their 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, stated:

“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.)

It appears, therefore, that reproducibility of assays among academic centers has not been established, such that it would not be feasible to compare an increase in Becker-type dystrophin of 0.2 to 0.3% (or even far greater increases) with dystrophin values cited in the literature for other mutations/patient populations, assessed by other laboratories.

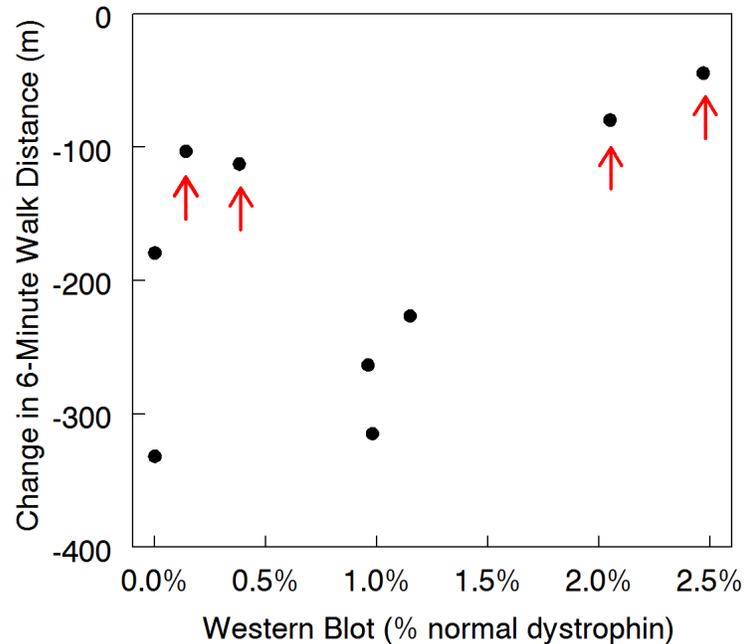
Do the clinical data bolster the question of “reasonably likely?”

The applicant collected data on both dystrophin production and physical performance in Study 201/202. Such data have the potential to support the concept that the dystrophin level predicts clinical response, and would support the ‘reasonably likely’ premise. Despite detailed testimonials from patients in Study 201/202 claiming improvements in clinical performance, the Division concluded, on the basis of the data presented in the NDA, that no patient in Study 201/202 clearly deviated from the natural history of the disease. They reasoned, therefore, that whatever the quantity of Becker-type dystrophin detected, it did not predict clinical benefit. Dr. Bastings opines that the clinical data weaken, and do not strengthen, the “reasonably likely” argument.

Within Study 201/202, it is also reasonable to consider the correlation between the quantity of dystrophin detected and maintenance of physical function in individual patients. The presence of a correlation would help support the “reasonably likely” question.

For the 9 patients who remained ambulatory at Week 180 and agreed to undergo a fourth muscle biopsy, Figure 10 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 4 patients whose 6-minute walk performance was best preserved (red arrows), 2 had the highest dystrophin levels detected in the study, but 2 had levels that were close to zero. Importantly, therefore, these data do not show a quantitative correlation between the surrogate endpoint deemed reasonably likely to predict clinical benefit, i.e., Becker-type dystrophin levels, and the clinical benefit, i.e., maintenance of walking velocity. In Dr. Bastings’ memorandum, he provides careful documentation of the trajectories of physical performance for each patient, comparing their changes in performance to the quantity of dystrophin detected. After careful consideration, he finds no correlation whatsoever.

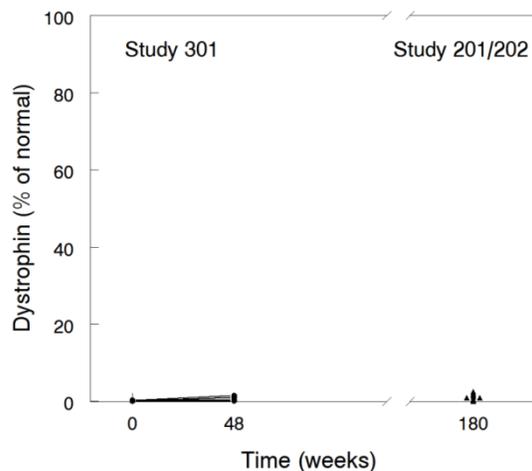
Figure 10: Study 201/202 – Lack of Correlation between Quantity of Dystrophin Detected and Preservation of Physical Function (6-Minute Walk Distance)



Although it should be obvious that changes on the order of a percent or two are small, it is nevertheless worthwhile to view these data at full scale to gain perspective (Figure 11). The figure is identical to Figure 9, except for the scale on the y-axis.

If dystrophin were simply an enzyme responsible for biochemical activity in myocytes, one could posit that a very small quantity of the protein could exert a substantial treatment effect, especially because levels are so low in untreated patients. But given that dystrophin is a structural support protein that helps prevent myocyte injury from 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. Finally, we must recognize receiving a treatment that increases dystrophin by 0.3% is not that same as being born with 0.3% more dystrophin.

Figure 11: Studies 301 and 201/202: Expression of Becker-type Dystrophin by Western Blot (full scale)



3. Dose-response

Although the issue is somewhat peripheral to the “reasonably likely” question, the presence of a dose-response in Study 201/202 would have provided supportive evidence that the dystrophin that was detected was produced by eteplirsen. A dose-response was not evident, although one could reasonably argue that the trial was very small and that the difference between 30 and 50 mg/kg/week was unimportant.

In a monkey study conducted to assess the pharmacodynamic effects of eteplirsen, a 1-log increase in dose (from 4 to 40 mg/kg) caused minimal increase in exon 51 splicing as detected by PCR (Section 4, Table 1). However, with a 2-log increase in dose (from 4 to 320 mg/kg), there was a log increase in exon 51 splicing. As noted in Section 4 of this memorandum, it is possible that much higher doses of eteplirsen could have a substantially greater effect, which might translate to clinical benefit.

Advisory Committee

The Advisory Committee was asked to discuss: a) the strength of evidence that eteplirsen increased the amount of dystrophin in muscles of treated patients relative to their baseline, and b) the clinical meaning of the amount of dystrophin observed in the muscles of eteplirsen-treated patients, taking into consideration the range of amounts of dystrophin known to be typically present in patients with DMD and in patients with Becker muscular dystrophy. (Of note, the data from Study 301 were not known/available to the Advisory Committee.)

Although the Committee failed to reach consensus on these questions, the discussion, summarized below, is of interest.

With respect to production of dystrophin, about half of the committee members found evidence that eteplirsen increased the amount of dystrophin produced in skeletal muscles. Among those who were not convinced, two members cited issues with the controls (lack of pre- and post-treatment biopsies in the same patients; differences in muscle groups biopsied), two had concerns about inconsistencies between dystrophin levels and clinical response (Figure 10), and one cited concerns about the lack of a dose-response (Table 3).

Only four Committee members had explicit comments with respect to the clinical meaningfulness of the amount of dystrophin detected in treated patients, and their opinions were split. One member opined that the amount of dystrophin needed to impart clinical benefit is unknown, but could be very low, or very low in a subset of patients. One of the patient representatives felt strongly that dystrophin was produced, and that the amount was sufficient to produce clinical benefit. One committee member, having opined that some dystrophin was produced, stated that there is no basis to determine the quantity of dystrophin that would be clinically significant, or whether the dystrophin is functionally active. Another committee member, one who had not opined on whether dystrophin was produced, noted that whatever the amount of dystrophin produced in the study, the amount was not clinically meaningful, based on the lack of correlation between dystrophin levels and clinical results (Figure 10).

The Committee voted on whether the applicant had provided substantial evidence from adequate and well controlled studies that eteplirsen induces production of dystrophin to a level that is reasonably likely to predict clinical benefit.

Ultimately, 7 members voted “no” and 6 voted “yes,” after one member changed his vote from “no” to “yes.” In explaining their “no” votes, 5 committee members opined that the studies were not adequate and well controlled; they questioned the techniques used to measure dystrophin as well as the appropriateness of the controls. Four committee members expressed concern about the lack of correlation between the dystrophin levels and clinical measures. They agreed that even if some dystrophin was produced, there was no evidence that dystrophin production was at a level that would be reasonably likely to predict clinical benefit. The 6 “Yes” votes included the consumer representative and 2 patient representatives. These individuals believed that there was some difference in dystrophin production and some evidence of improvement in endpoints. One of the members who voted “Yes” stated that he was very troubled by not understanding what constitutes a clinically significant amount, but was impressed by the patients’ observations. Two members who voted “No” stated that their vote was justified by the way the question was phrased, but that the patient testimonies suggested the drug works.

Is There a Basis for a Conventional Approval Based on Clinical Data?

The clinical data have been well described by the review team. The development program consisted of one trial (Study 201/202) with a relatively short (24-week) placebo-controlled portion (Study 201) followed by a long-term extension study (Study 202). Although the applicant submitted biopsy data from the ongoing Study 301, no clinical data have been submitted from that study.

As noted above, for Study 201, patients were randomized (1:1:1) to eteplirsen 30 mg/kg/week, eteplirsen 50 mg/kg/week, or placebo (n=4 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) and followed for 4 additional weeks. The trial was extended to an open-label phase (Study 202), where all 12 patients continued to receive eteplirsen without interruption, although investigators and patients remained blinded to dose.

The 1° endpoint of Study 201 was the percentage of dystrophin-positive fibers in muscle biopsies as assessed using immunohistochemistry, but there were numerous exploratory endpoints.

When the data from Study 201 were originally analyzed, the applicant found that eteplirsen caused a striking and unprecedented increase in dystrophin production, based on the reading of the immunohistochemistry data at Nationwide Children’s Hospital, with supportive data from Western blot analyses.

The clinical data, too, were interpreted as positive. As discussed by the review team, 2 patients in the 30 mg/kg/week treatment group became unable to ambulate soon after the trial began, and there were no significant differences in 6-minute walk distance among the groups. Despite clearly negative results, the applicant performed a *post hoc* analysis that omitted the 2 patients in the eteplirsen group who became unable to ambulate. They represented these results as positive, and publically promoted both the immunohistochemical dystrophin results and the 6-minute walk data as positive (see clinical review).

Although FDA would later determine that the analyses underlying these data were not valid, the publicity from the paper² and Sarepta’s press release³ raised unrealistic expectations of efficacy

in the DMD community. It was these perceptions that led the applicant to conclude that a second placebo-controlled study would not be feasible.

FDA strongly suggested a second, larger, adequately-powered, placebo-controlled trial, but the applicant was reluctant to run such a trial, in part because their supply of drug was limited, and in part because of their insistence that the DMD community would not agree to participate in a trial where there was a chance of receiving placebo. Faced with the applicant's unwillingness to conduct a second placebo-controlled trial, FDA agreed to an externally-controlled trial: a comparison between patients in the ongoing Study 202 and patients in an external control group. The Division expressed strong concern, however, with respect to the interpretability of such a trial with 6-minute walk distance as the endpoint, given that physical performance is not a "hard" endpoint, but can be influenced by motivation and other factors. Citing FDA Guidance,⁵ the Division noted that the treatment effect would have to be dramatic for the results from an externally-controlled study to be interpretable. Details of the interactions between FDA and Sarepta are well documented by the review team.

International guidelines, adopted by the FDA as guidance, stress caution with respect to the interpretation of data from externally-controlled trials. As noted in the International Conference on Harmonization (ICH) E10 Guideline,⁵ blinding and randomization, used to decrease bias in randomized controlled trials, are not utilized in externally-controlled trials; the inability to control bias is a critical limitation of externally controlled trials. Groups can be dissimilar with respect to a wide variety of factors that could influence outcome – factors that are both known and measurable as well as factors that are unknown. As explained by Dr. Robert Temple at the April 25, 2016 meeting of the Peripheral and Central Nervous System Drugs Advisory Committee, it has been well documented that untreated historical-control groups tend to have worse outcomes than apparently similarly chosen control groups of randomized studies, possibly reflecting a selection bias.

The ICH E10 Guideline explains: "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 inability to control bias restricts use of the external control design to situations in which the effect of treatment is dramatic and the usual course of the disease highly predictable. In addition, use of external controls should be limited to cases in which the endpoints are objective and the impact of baseline and treatment variables on the endpoint is well characterized." In essence, in order to be interpretable, the finding of a difference between groups should be large – so large that the difference is patently obvious without the need to rely on inferential statistics.

Having heard FDA's concerns regarding the potential difficulty in interpreting an externally-controlled trial, the applicant nevertheless obtained access to individual data from patients with DMD from Professor Eugenio Mercuri at the Catholic University in Rome on behalf of the Italian DMD Registry database (n=97) and from Professor Nathalie Goemans at the University Hospitals in Leuven (n=89). From these 186 patients, 50 had a genotype amenable to exon skipping therapy, were using corticosteroids at baseline, had 6-minute walk data available at baseline, and were ≥ 7 years old. Among these 50 patients, 13 had a genotype amenable to

⁵ Guidance for Industry: E10 Choice of Control Group and Related Issues in Clinical Trials, May 2001

exon 51 skipping therapy. I will note that the review team has been unable to gain an understanding of how dates of inception were determined for registry patients, i.e., when patients were considered to have ‘enrolled.’

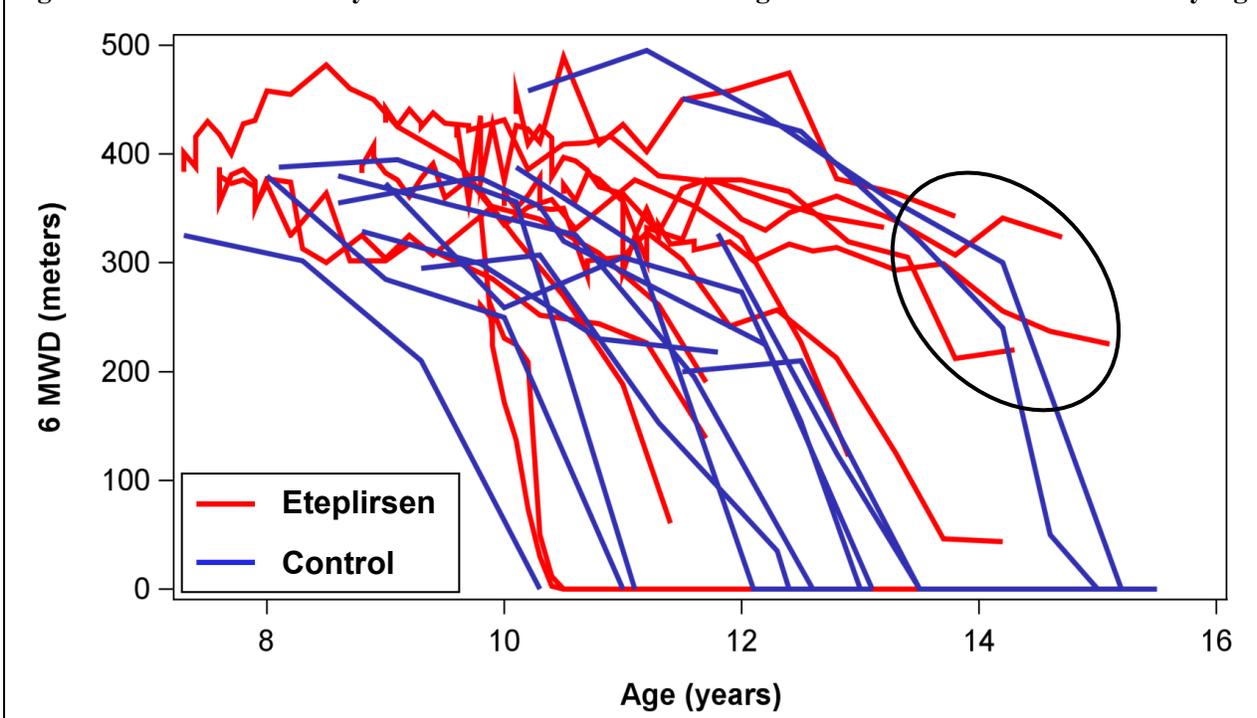
Study 202 was continued, therefore, with patients continuing to receive either 30 or 50 mg/kg/week eteplirsen. Numerous comparisons of physical function were planned between these 12 patients and the 13 patients in the external control group. Measures included 6-minute walk, rise time, timed 10-meter run, and North Star Ambulatory Assessment (NSAA).

With two small groups of patients, there was no way to match patient pairs. Fortuitously, the mean ages and 6-minute walk distances were well matched at baseline, although the review team found that initial age of steroid use and baseline NSAA scores were dissimilar between groups – and both of these differences favored the eteplirsen group.

It is clear that some patients exited the registry to enroll in clinical trials. Thus, DMD patients who remained in the Italian and Belgian registries (the control group): 1) did not seek knowledge (or lacked knowledge) regarding applicable clinical trials into which they might have enrolled; 2) sought enrollment in trials but did not qualify; or 3) qualified for enrollment in a trial(s) but made a conscious decision not to participate. Obviously, such patients could differ substantially from patients in Study 201/202. The point is that there can be unknown factors beyond baseline age, weight, length of steroid use, and 6-minute walk distance that importantly affect outcomes.

The applicant presented the data by time-on-treatment, but because physical abilities change significantly with age in patients with DMD, the review team believes that the more meaningful way to display the longitudinal 6-minute walk data is by age (recognizing that both analyses have advantages and limitations, and that there is no ideal way to present these data). The 6-minute walk data are shown in Figure 12 as a function of age. The review team stresses that,

Figure 12: Patients in Study 202 vs. Patients in External Registries: 6-Minute Walk Distance by Age



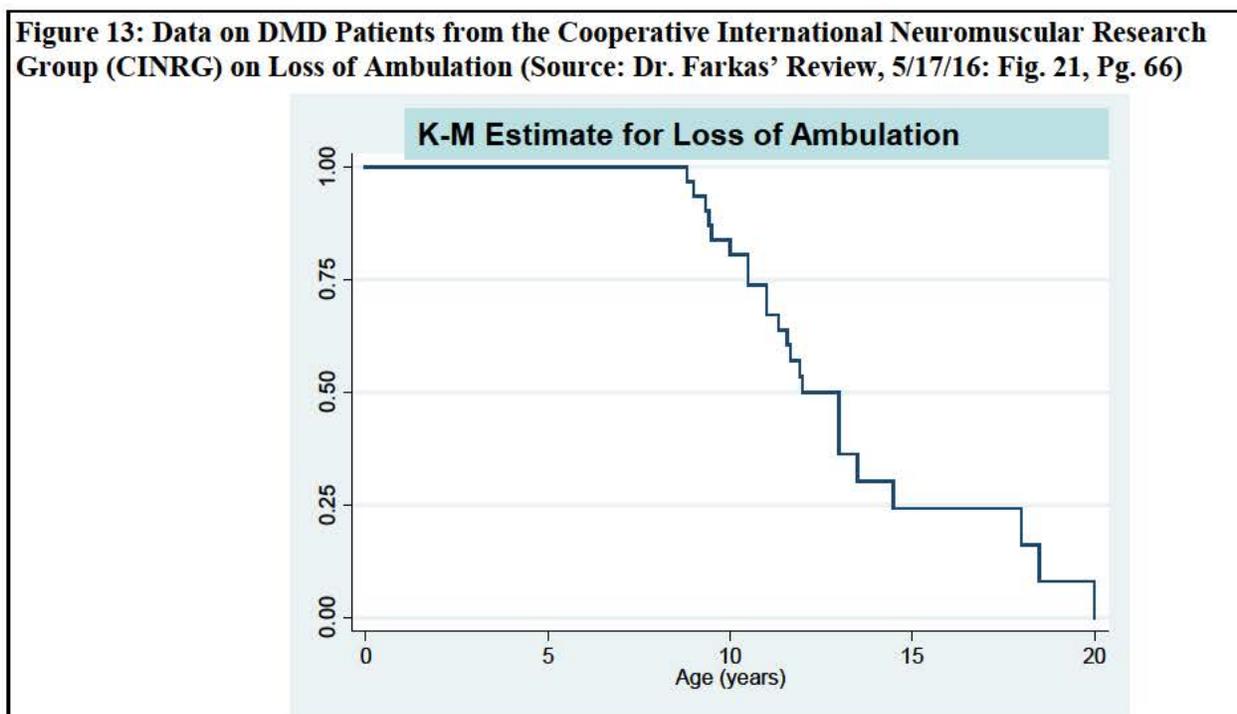
by simple visual inspection, the two groups show little difference in performance.

There are 4 patients in the eteplirsen group, ~14 to 15 years of age, who continue to retain good walking ability (inside the oval). There are 2 control patients in this age range who had been maintaining similar walking ability, but appear to have experienced a precipitous loss of ambulation between ages 14 and 15. As explained by the review team, there are concerns regarding the comparability of the assessments of these patients, and concerns about comparability of the groups in general.

The applicant's argument for accelerated approval is based on this comparison of 6-minute walk distance between the patients in Study 202 and the patients in the external control group from Italy and Belgium. The difference in 6-minute walk distance is certainly statistically significant. The problem is that the study was externally-controlled, and the statistical test was based on a non-randomized comparison.

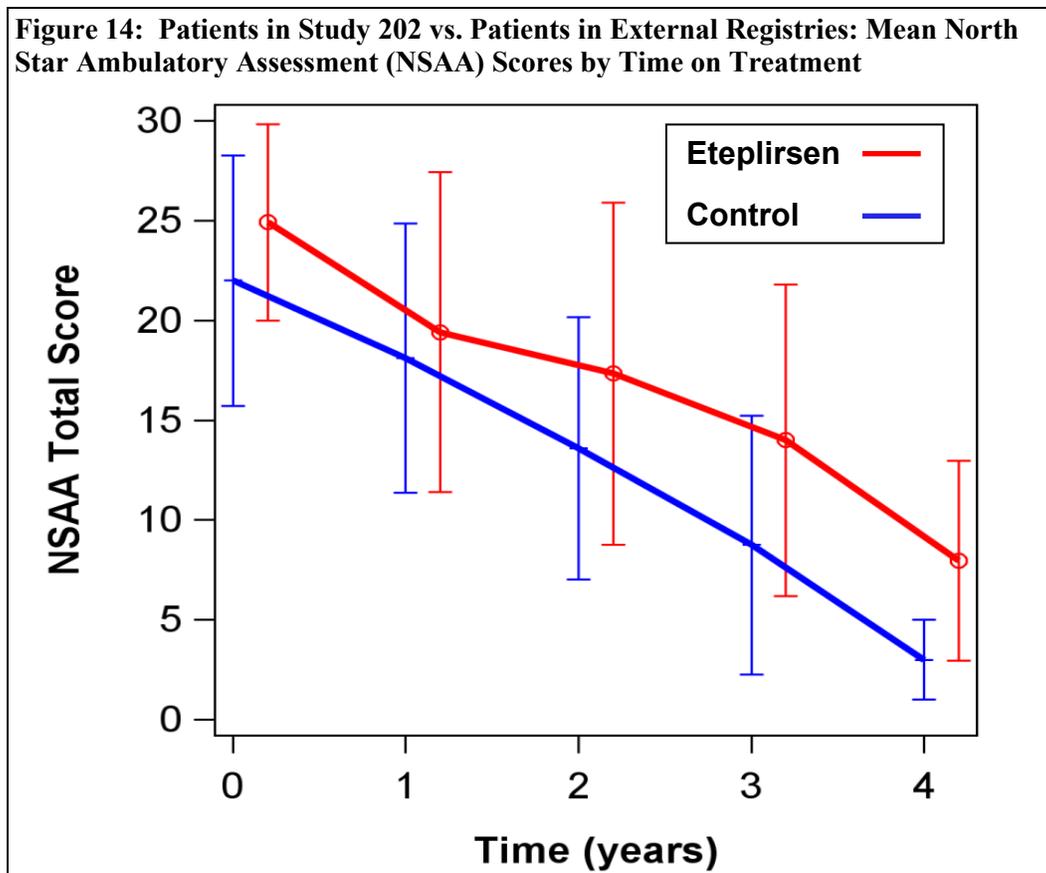
Data from the Cooperative International Neuromuscular Research Group (CINRG) provide an additional source of information on the natural history of patients with DMD. Figure 13 is a Kaplan-Meier (K-M) survival curve from CINRG showing time-to-loss of ambulation. Of note, 25% of patients remain ambulatory at age 17; their course seems quite consistent with that of patients from Study 201/202.

Figure 13: Data on DMD Patients from the Cooperative International Neuromuscular Research Group (CINRG) on Loss of Ambulation (Source: Dr. Farkas' Review, 5/17/16: Fig. 21, Pg. 66)



In summary, the review team strongly believes that patients on eteplirsen in Study 201/202 do not demonstrate a substantial treatment effect on walking velocity that clearly differentiates their course from the natural history of the disease. For a more complete description with comprehensive patient profiles, see the reviews of Drs. Breder and Farkas and the memo of Dr. Bastings.

Finally, as stressed by the review team, the data from other measures of physical function, i.e., rise time, timed 10-meter run, and North Star Ambulatory Assessment (NSAA), show steady decline in the eteplirsen-treated patients that does not differ substantially from the decline in the external control group. The NSAA data are shown in Figure 14 by time on treatment (eteplirsen patients) or time since inception (registry patients). The NSAA is thought to be a comprehensive outcome measure, well reflecting the functional abilities of DMD patients. Of note, the downward trajectories of the two groups are indistinguishable (the lines are virtually parallel with equal slopes).



Patient Testimony/Advisory Committee:

In addition to the presentations made by the applicant and the review team at the April 25, 2016, Advisory Committee Meeting, there were testimonies from over 50 individuals and families, including most of the patients who were participating in Study 202. (Per email communication from (b) (4), one of the applicant’s consultants, 10 of the 12 patients testified and another patient had someone speak on his behalf.)

In addition, the applicant invited Christine McSherry, Executive Director of the Jett Foundation, to present “Patient and Caregiver Reported Outcomes of Patients in Clinical Trials of Eteplirsen for Treatment of Duchenne.”

The testimonies of these patients were quite consistent and remarkably positive: all were convinced that eteplirsen had made a substantial positive impact on their physical performance, improving numerous aspects of their lives.

It was noteworthy that a number of individuals who were in Study 201/202 reported *improvement* in physical function with eteplirsen treatment. For example, one patient stated that he had required a wheelchair at a school he had attended in the past, whereas he no longer needed a wheelchair at his present school. A video showed a boy who, prior to treatment, had some difficulty climbing up into the seat of a minivan. After receiving eteplirsen for several months, he was shown jumping up easily into the seat. In another video, a boy in the study threw a football, a tight spiral, with ease and finesse.

Many of the Committee members seemed obviously moved and deeply affected by these testimonies and videos, seemingly convinced that there was a treatment effect.

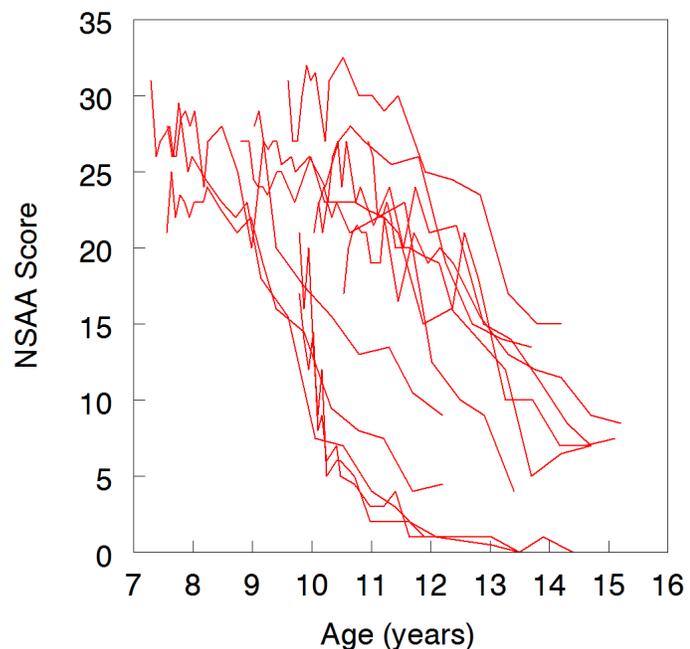
Importantly however, despite the claims of improvement made at the microphone at the Advisory Committee meeting, the review team did not find any patients in Study 201/202 with consistent improvement in physical performance as assessed by formal testing (6-minute walk, rise time, NSAA, 10-meter run). These tests have shown moderate to extreme declines in physical function for all patients (see NSAA data, Figure 15).

Thus, the review team and many on the Advisory Committee (including Benjamin Dupree, the patient representative with DMD), were unable to reconcile the patient testimonies with the data collected by the applicant: the testimonies spoke of *improvement*; the data showed *progressive worsening*.

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.

The 7-member majority of the committee who voted "no" agreed that Study 201/202 was not a well-controlled study. Most cited problems with the controls. One member explained that a historically-controlled study *could* provide evidence of effectiveness, but that Study 202 did not. Two committee members noted that the original placebo-controlled portion of the study was

Figure 15: Study 201/202 – Individual NSAA Performance by Age



negative. One member who cited issues with the controls also noted that a single trial would be insufficient to provide substantial evidence.

The 3 members who voted that there was substantial evidence of effectiveness explained that the study results correlated with the testimonies presented by the public.

Commentary:

I agree with the Division, the Office of Biometrics, the Office of Clinical Pharmacology, and the Advisory Committee with respect to the lack of substantial evidence of effectiveness for eteplirsen. The review team elaborates on many factors that differ, or could differ, between the treatment groups – factors that could lead to a difference in outcomes. Externally-controlled trials are best-suited for diseases where progression is highly predictable and treatment effects are extreme. Although there appeared to be a difference in ambulation between patients in Study 202 and patients in the external control group, the effect size was not sufficient to be persuasive, given the inability to control bias in an externally-controlled study. As explained in ICH E10, "...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." With only 12 patients in the trial and a moderate difference in walking velocity, the study falls short.

Finally, it is critical to note that no dose-limiting side effects were observed at either dose tested in Study 201/202, and even the most optimistic interpretation of the data is that patients experienced gradual decline in function – not stabilization. Even if one were to reach the conclusion that the applicant showed substantial evidence of dystrophin production, deserving of accelerated approval, investigation of higher doses would be imperative.

8. Safety

As explained in the clinical review, the number of subjects exposed was too small to provide an adequate assessment of safety. On the other hand, I also agree with the review team that the deficiencies in safety assessments would not likely be an issue for approvability in their own right had the drug been demonstrated to be effective. In other words, for a therapy that is shown to be effective in a serious condition where there are no approved drugs, we would approve a marketing application even with substantial risks, as long as we could write adequate instructions for use. Moreover, we would not delay approval of a marketing application because of uncertainty of risks. Instead, we would work with the applicant to obtain more extensive safety data post-approval. Such would be the case for this application if there were substantial evidence of effectiveness.

Of note, many patients in these studies are now receiving infusions through chronic indwelling catheters. Although we are not aware of any serious adverse events caused by infections, with approval of this drug there would undoubtedly be serious infections and possibly rare deaths eventually. The risk of an indwelling IV line in patients on chronic corticosteroids should be mentioned in labeling if the drug is approved.

Although neither immunogenicity nor allergic reactions have been reported with eteplirsen, immunogenicity testing would be advisable in ongoing trials. Moreover, given that these

patients may be naïve to Becker-type dystrophin, the potential for anti-dystrophin antibodies should be studied as well.

9. Advisory Committee Meeting

There were many important discussions at the April 25, 2016 Advisory Committee Meeting, and they are summarized above, in context.

10. Pediatrics

Duchenne Muscular Dystrophy is an orphan indication, not subject to the Pediatric Research Equity Act.

11. Other Relevant Regulatory Issues

Site Inspections:

The site at Nationwide Children’s Hospital was inspected in 2014. See description and conclusions in Section 7, above, and, in particular, the summation and discussion in Dr. Breder’s review.

Dr. Ashutosh Rao conducted an inspection with Young Moon Choi, Ph.D. (Office of Study Integrity and Surveillance) and Mark Babbit (Office of Regulatory Affairs) of the facilities at University of Iowa in Iowa City, IA and Sarepta Therapeutics Inc. in Corvallis, OR. The inspections confirmed that the blinding procedure, handling of the sample shipment, and the conduct of Western blot analyses of the samples from Study 301 (PROMOVI) were consistent as predefined in the protocol.

Name Review:

The Division of Medication Error Prevention and Analysis concluded that the proposed proprietary name, “EXONDYS 51,” is acceptable from both a promotional and safety perspective.

12. Labeling

I do not recommend approval, but if the drug were to be approved, the label would need to state that no clinical benefit has been established, and explain the effect on the surrogate endpoint in clearly understandable language (i.e., 0.3% or 3 parts in a thousand). Section 6 would need to note that safety is not well characterized.

13. Decision/Action

DMD is a rare genetic disease characterized by the near absence of functional dystrophin protein, leading inexorably to myocyte degeneration, muscle dysfunction and inflammation, severe disability, and death, robbing patients of their dignity along the way. Although steroids are thought to slow the course of the disease and are typically considered standard of care, they are by no means curative, and they have their own side effects.

The cause of DMD is well established – the absence of structural dystrophin protein in myocytes. There is wide belief in the medical/scientific community that restoration of functional dystrophin protein has a strong potential to ameliorate the disease.

Eteplirsen is a novel PMO that is designed to lead to translation of an abnormal but functional dystrophin protein – a protein that is produced in Becker muscular dystrophy, a far less severe form of muscular dystrophy. The data from RT-PCR show that the drug produces the intended Becker-type messenger RNA; we have no data on the extent of messenger RNA production.

As noted by the review team, the clinical data generated from study 201/202 do not provide evidence of efficacy. The aim of Study 201, the only randomized placebo-controlled study conducted by the applicant, was to assess dystrophin production in response to lower and higher eteplirsen regimens (30 or 50 mg/kg/week) vs. placebo. Results of the original analyses of Study 201, published in a major journal, were remarkably positive, and their publication led to widespread enthusiasm for the drug. Unfortunately, an FDA inspection found a number of important technical factors that rendered the data unreliable and uninterpretable: the Western blot analyses were sub-standard; there were also critical problems with the reading of the immunohistochemistry images. FDA recommended a blinded re-read of the images, but upon re-read of the images by 3 blinded pathologists using FDA-recommended procedures, there was no increase in dystrophin production.

Likewise, Study 201 did not meet its 1^o clinical endpoint, 6MWT, at Week 24. Two patients in the low-dose eteplirsen group became unable to ambulate early in the study, such that a proper intent-to-treat analysis of the 6-minute walk data nearly showed a statistically significant difference in favor of *placebo*.

The applicant switched all patients to active drug in Study 202, and has continued to follow the patients for 6-minute walk distance, NSAA, and rise time.

Study 202 did not meet its 1^o clinical endpoint, 6MWT, at 48 weeks.

The alternative analyses of Study 202 proposed by the applicant are based on comparison to an external control group obtained from registry patients in Italy and Belgium. Questions about comparability notwithstanding, analyses have not shown a clear separation of the disease course between eteplirsen-treated patients and external controls. Moreover, there is not a clear separation between eteplirsen-treated patients and patients in the CINRG registry. Thus, neither external control group suggests there is a treatment effect.

The Western blot analyses from Week 180 of Study 201/202 showed a low quantity (0.9%) of dystrophin; however, the study was not adequate and well controlled (the baseline level of dystrophin was not known with certainty), and the lack of correlation between results of Western blot and immunohistochemistry demonstrates a troubling lack of internal consistency.

Study 301, on the other hand, was an adequate and well-controlled study that provided substantial evidence of Becker-type dystrophin production in response to eteplirsen. The mean change in Becker-type dystrophin with treatment was 0.22% to 0.32%, depending on the method used to impute values less than the lower limit of quantification. Although all members of the review team believe that Becker-type dystrophin is an appropriate surrogate endpoint, the mean quantity of dystrophin produced in Study 301 was minute by any standard. In considering

responders, even the largest responder in Study 301 produced only 1.33% of normal dystrophin, which is thought by many authorities to be insufficient. No other patient produced 1% dystrophin in response to treatment.

Recognizing that the threshold for the effect size needed to be 'reasonably likely' to predict clinical benefit is not known, the view provided in the literature suggests that at least 3% of normal dystrophin is inadequate, and levels perhaps much more, a minimum of 10%, would be necessary for detectable clinical benefit. The finding in Study 301, an increase in the range of 0.22 to 0.32% of normal, is an order of magnitude below this level.

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 be capable of ameliorating the fundamental genetic defect of DMD, but the effect size seems insufficient at the tested doses.

Various individuals have opined that there appears to be some evidence that some patients are producing dystrophin in response to eteplirsen; however, such optimism fails to reach the legal threshold of 'reasonably likely to predict clinical benefit' required for accelerated approval.

Accelerated approval of this NDA based primarily on the change in Becker-type dystrophin in Study 301 would be problematic for these reasons:

1. The amount of dystrophin produced in Study 301 is so meager that it could be considered to be tantamount to any increase in dystrophin. In other words, if a statistically significant change of 0.22% is considered adequate to support accelerated approval, then the question arises as to whether there is any statistically significant change that would be too small to support accelerated approval. Similarly, if a response had been defined as a treatment effect of 1%, there would have been only one (out of 12) responders in Study 301.

If we were to adopt the concept that, for rare diseases, accelerated approval can be supported by any statistically significant change in an appropriate surrogate (or by a response in a single patient), we would enable accelerated approval of numerous drugs for rare diseases. No doubt there are some who would applaud this as a regulatory advance, but these are typically the kinds of findings that support Breakthrough Designation, not approval. If accelerated approval based on any change in a surrogate endpoint is what is meant by regulatory flexibility and this is the new normal, a new approval pathway is clearly needed.

With lowering of the standard for accelerated approval, the result would be a world where traditional clinical trials are abandoned in favor of small proof-of-concept studies designed to show any level of production of a target protein – e.g., a statistically significant effect in a paired pre- vs. post-treatment analysis that is clinically meaningless. There would be no reason to pursue placebo-controlled clinical trials to support efficacy prior to accelerated approval; in fact, the possibility of failure would provide a substantial disincentive to the conduct of such trials. 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.

2. Even if the 30 mg/kg/week dose were considered to have a meaningful effect on the surrogate endpoint, the dose is sub-therapeutic. Moreover, the short 3.5-hour half-life of eteplirsen by no means supports a weekly dosing regimen. 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.

Imagine that 100 years ago a promising drug called penicillin is discovered – a potential cure for pneumococcal pneumonia – but the drug is difficult to produce and expensive. A dose of 5 mg weekly has been shown to have statistically significant bactericidal effects on *Streptococcus pneumoniae*. Would it be ethical to give the drug accelerated approval based on this finding and allow marketing of a dose of 5 mg, absent additional information? (The therapeutic dose is ~2 logs higher than 5 mg.) Patients who might receive a lifesaving therapy (i.e., a higher dose) would die because the dose is too low.

Despite considerable pressure from the DMD patient community and many well-intentioned members of the public who have lobbied on their behalf, I am unable to reach the conclusion that the applicant has provided substantial evidence to support either conventional or accelerated approval of eteplirsen for the treatment of DMD. This view is in agreement with the unanimous opinions of members of the review team from the Division of Neurology Products, the clinical pharmacology review team, and the biostatistics review team. The Advisory Committee was under intense and near-incessant pressure from a large public audience, urging them to believe that eteplirsen was effective, and life changing in some circumstances. Emotions in the room ran high. In spite of this pressure, that majority of the Advisory Committee voted against both conventional and accelerated approval.

In a June 3, 2016 letter from Dr. Janet Woodcock, the applicant was advised that “If you are successful in showing, to FDA’s satisfaction, a meaningful increase in dystrophin by Western blot analysis between the paired pre-and post-treatment samples, we expect to be able to grant an accelerated approval....” It is difficult to consider production of 2 to 3 parts per thousand as a “meaningful” change. To put this effect into perspective, if a normal amount of dystrophin were equivalent to a \$5 bill, this change would be equivalent to a penny.

With all of this information at hand, most sponsors would have concluded that exploration of higher doses was needed; however, this applicant chose instead to trumpet the preliminary findings from their 12-patient phase 1/2 study, convincing many in the DMD community that the drug was highly effective, and unleashing a public media campaign (with support of many politicians) to approve the drug. The reality is that FDA is a science-based organization. We do not – and should not – make approval decisions based on patient anecdotes or campaigns through social media.

I strongly agree with the decisions of Dr. Bastings, reviewer staff in the Division, the Office of Biometrics, and the Office of Clinical Pharmacology to issue a complete response for this NDA. I also agree that it would be desirable to provide access to this drug for DMD patients through expanded access programs, with cost recovery, while an adequate dose-finding study is conducted.

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 a dose of 30 and 50 mg/kg/week are adequate, but they do not support efficacy.

We 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 1° endpoint, again, testing higher doses.

The applicant is continuing to enroll the PROMOVI study, an open-label, multi-center, 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.

The 1° endpoint is change in 6-minute walk test distance from baseline. A 2° endpoint is the percentage of dystrophin-positive fibers, as assessed by immunohistochemistry. Patients undergo muscle biopsies at baseline and various time points to assess dystrophin production.

My suggestion for a path to approval is to randomize patients in the ongoing PROMOVI study to:

- 1) 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 blind to treatment group. For accelerated approval, the 1° endpoint would be dystrophin production, comparing the higher and lower doses. For standard approval, the 1° endpoint would be a test(s) of physical performance such as rise time or the NSAA.

Such a trial would be methodologically sound and ethical. Virtually everyone, patients and physicians alike, want to know if higher eteplirsen doses would increase dystrophin production, and would have equipoise for participation. Although there is concern regarding performance of muscle biopsies in patients assigned to placebo, this concern would not exist in this study. And if the applicant were to forego immunohistochemistry studies, needle biopsies with local anesthesia (rather than open biopsies under more intensive anesthesia) would be sufficient.

This study design would simultaneously address another concern that I believe has been underappreciated by many. As noted above, it would be problematic in my view to approve a dose of 30 mg/kg/week, presumably leading to a dystrophin increase of ~0.3%, when it is

known that this dose fails to prevent the decline in physical function and yet produces no overt toxicity. The monkey data (Table 1) suggest that much higher doses might have a far greater effect on exon skipping, an impact that might prevent disease progression. Thus, it seems imperative to study higher exposures.

14. Final

Many of us would wish to approve this drug if we could. DMD is a horrible disease and there are no approved treatments. FDA takes seriously the patient perspective and our congressional mandate to be flexible. But patient-focused drug development is about listening to patient perspectives about what matters to them; it is not about basing drug approvals on anecdotal testimony that is not corroborated by data.

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.

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

/s/

ELLIS F UNGER
07/16/2016