

**CENTER FOR DRUG EVALUATION AND
RESEARCH**

APPLICATION NUMBER:

206488Orig1s000

SUMMARY REVIEW



TO: Janet Woodcock, M.D., Director, CDER
Ellis Unger, M.D., Director, Office of Drug Evaluation I, CDER
Luciana Borio, M.D., Chair, Agency Scientific Dispute Process Review Board

FROM: Robert M. Califf, M.D., Commissioner of Food and Drugs

RE: Scientific Dispute Regarding Accelerated Approval of Sarepta Therapeutics' Eteplirsen (NDA 206488) – **Commissioner's Decision**

DATE: September 16, 2016

Overview

The Chair of the Agency Scientific Dispute Process Review Board (SDR Board), Dr. Luciana Borio, has forwarded to me the Board's recommendation regarding the appeal by Dr. Ellis F. Unger, the Director of the Office of Drug Evaluation I, of Dr. Janet Woodcock's decision; namely, that Sarepta Therapeutics' drug eteplirsen (AVI-4658) meets the standard for accelerated approval for treatment of Duchenne muscular dystrophy (Duchenne or DMD) in patients who have a confirmed mutation of the dystrophin gene amenable to exon 51 skipping, which affects about 13 percent of the population with DMD. Duchenne is a rare genetic disorder characterized by progressive muscle deterioration and weakness and is often fatal. Dr. Unger disagrees with Dr. Woodcock's view and would decline to approve eteplirsen at this juncture.¹

The SDR Board's memorandum of August 8, 2016 concludes that under the standard for review set out in FDA's Staff Manual Guide for formal appeals of internal FDA scientific disputes, Dr. Unger had an adequate opportunity to present his scientific concerns within the Center for Drug Evaluation and Research (CDER), and that Dr. Woodcock considered all relevant evidence in making her decision.² In her capacity as Acting Chief Scientist, Dr. Borio also conveyed her own views on the dispute, stating among other things that she does not believe the available data and information support accelerated approval of eteplirsen.³ As the SDR Board's role was to conduct a procedural review, the SDR Board further recommended that I either conduct a substantive scientific review of the dispute or convene a panel of experts to conduct a scientific review and

¹ Appeal at 1.

² SDR Board Recommendation at 2.

³ *Id.* at 2, 25-26.

advise the Agency on whether the available evidence for eteplirsen meets the standard for accelerated approval.⁴

I note that, in my understanding, it is highly unusual for a Center Director's decision regarding a product application to be appealed to the Commissioner's office. Decisions on medical product approvals and clearances are delegated under provisions of Staff Manual Guide 1410 to the appropriate Center Directors and Center staff.⁵

I have read the documents pertinent to Dr. Unger's appeal, Dr. Woodcock's final Center decision, and the review and recommendation of the SDR Board. I have performed a thorough review of the basic and clinical science, a review that also comprised a briefing with the review team, including Dr. Unger, and discussions with Drs. Woodcock and Jenkins. Although under the standard for review set forth in Staff Manual Guide 9010.1, I was not required to review the science, I did so in order to properly evaluate the positions of Dr. Woodcock and Unger. **My decision following this review is to defer to Dr. Woodcock's judgment and authority to make the decision to approve eteplirsen under the accelerated approval pathway, in her capacity as Director of the Center for Drug Evaluation and Research.**

My reasoning and detailed considerations for my decision are given below.

Examination of the Scientific Dispute

Key Points of Agreement

There is agreement that Dr. Unger's views and those of the review team were heard in great detail in an environment of open discussion and dissent on multiple occasions. I agree with the SDR Board that there is no basis for overturning Dr. Woodcock's decision on procedural grounds and Dr. Unger also agrees with this conclusion.

In addition, the science is not in dispute beyond the usual types of disagreement that occur when experts review clinical evidence from different perspectives. As discussed in detail below, the evidence evaluation is complicated by the unusual circumstance of the development program, which involved a small subset of children with a rare disease and was characterized by major

⁴ *Id.* at 2, 27.

⁵ Although, as described in Staff Manual Guide 1410.21.1.A., the Commissioner retains authority to make decisions on medical product applications, those decisions are most appropriately vested in the Centers and only the most unusual circumstances would warrant overturning the decision of a Center—including that of a Center Director—on a medical product application. As is set out in the Staff Manual Guide 9010.1 applicable to this matter, FDA has adopted a formal scientific dispute resolution process that emphasizes 1) establishing and maintaining robust dispute resolution processes within the Centers to ensure candid and comprehensive discussion of scientific issues, and 2) a Commissioner's Office review that focuses on the adequacy of the internal Center scientific dispute process (i.e., whether the Center's processes were followed, whether the Center considered all relevant evidence bearing on the scientific question at issue, and whether the initiator of the appeal was provided an opportunity to express concerns at all appropriate levels, prior to and including the Center Director). This established dispute resolution process does not contemplate a substantive decision on the disputed issue by the Commissioner.

flaws in the clinical study design, making the judgment on science difficult. The points of agreement include the following:

- DMD is a rare, progressive disease characterized by the virtual absence of functional dystrophin.⁶
- There is no approved therapy for DMD.⁷
- The principal question in this dispute is not whether dystrophin level is an appropriate surrogate endpoint for the purpose of accelerated approval, but whether the quantity of dystrophin produced by eteplirsen is an effect that is reasonably likely to predict clinical benefit, as required under section 506(c)(1)(A) of the Federal Food, Drug, and Cosmetic Act and FDA regulations at 21 CFR 314.510.⁸
- Available preclinical work supports the clinical development of eteplirsen, most notably clear documentation of production of the transcript and the protein in a dose-dependent manner in primates, without measurable toxicity.⁹
- Major flaws in both the design and conduct of the clinical trials using eteplirsen have made it impossible to use much of the resulting trial data as reliable evidence in regulatory decision-making, including for reasonable extrapolation to clinical care.¹⁰
- Despite the flaws in the clinical development program noted above, both Dr. Woodcock and Dr. Unger, as well as the review team, agree that eteplirsen produces measurable increases in dystrophin compared with control.¹¹ All agree that the amount of dystrophin produced is small compared with expectations at the outset of trials in humans.¹² Although the gap between expectations and measured increase is not directly relevant to the final decision, the log-order difference has added further uncertainty. If the levels of dystrophin had reached those measured in the setting of Becker muscular dystrophy, there would be much less concern.
- Although there is some disagreement about whether the data support the expression of dystrophin in one study (201/202), there is agreement that the other (study 301) demonstrates clear evidence for dystrophin production. Dr. Unger concludes that there is

⁶ Appeal at 2; SDR Board Recommendation at 2.

⁷ Appeal at 2. I note that there is a modest body of evidence suggesting that treatment with corticosteroids yields relatively short-term improvements in function and muscle strength (see Matthews E, et al. *Cochrane Database Syst Rev.* 2016;(5):CD003725. doi: 10.1002/14651858.CD003725.pub4), but there is no evidence from adequate and well-controlled trials that this therapy has a major effect on the course of the disease.

⁸ Appeal at 3; SDR Board Recommendation at 4, 17.

⁹ ODEI Decisional Memo, July 2016.

¹⁰ Woodcock Decisional Memorandum at 2-3, 11; Appeal at 4-5, 8-9.

¹¹ Woodcock Decisional Memorandum at 3; Appeal at 6.

¹² Woodcock Decisional Memorandum at 11.

evidence from one adequate and well-controlled trial, while Dr. Woodcock concludes there is evidence from two adequate and well-controlled trials.¹³

Key Points of Disagreement

There is disagreement about whether the amount of dystrophin protein produced is sufficient to be “reasonably likely” to predict a clinical benefit. The basis for the judgment about “reasonably likely” is subjective and is not explicitly defined in statute or regulations. FDA’s Guidance on Expedited Programs for Serious Conditions – Drugs and Biologics (May 2014) states at page 19 (internal citation omitted):

Determining whether an endpoint is reasonably likely to predict clinical benefit is a matter of judgment that will depend on the biological plausibility of the relationship between the disease, the endpoint, and the desired effect and the empirical evidence to support that relationship. The empirical evidence may include “. . . epidemiological, pathophysiological, therapeutic, pharmacologic, or other evidence developed using biomarkers, for example, or other scientific methods or tools.”

I conclude that Dr. Unger and Dr. Woodcock have each exercised reasonable scientific judgment in reaching differing conclusions on whether the effect on dystrophin production seen in the studies in Sarepta’s application reasonably predicts clinical benefit for the relevant subpopulation of Duchenne patients. There is also abundant evidence that Dr. Woodcock heard and read all the scientific evidence, including the detailed views of Dr. Unger and the review team; yet she came to a different conclusion. These differing conclusions are detailed immediately below:

- Dr. Unger, the review team, and Dr. Borio have concluded that the demonstrated levels of dystrophin are not “reasonably likely” to predict clinical benefit based on a chain of logic that is well-described in Dr. Unger’s appeal and in Dr. Borio’s summary.¹⁴ In essence, based on the belief that Becker muscular dystrophy provides a model for Duchenne because the dystrophin is similar to that produced with eteplirsen, Dr. Unger concludes that functional protein in the range of 10% of normal levels would be needed to provide evidence that would make it “reasonably likely” that changes in the surrogate would predict clinical benefit.¹⁵ The use of Becker muscular dystrophy as a model for DMD is an extrapolation, but a rational one in the absence of validation of dystrophin as a surrogate.
- Dr. Unger also finds that the results of the eteplirsen clinical trials to be inadequate for supporting the conclusion that there is any relationship between measured levels of dystrophin and improvement in function within the trial data sets. Making certain

¹³ Woodcock Decisional Memorandum at 2-5; Appeal at 4-6.

¹⁴ Appeal at 4, 11-20; SDR Board Recommendation at 25.

¹⁵ Appeal at 12-15.

assumptions, he provides plots and calculations that he interprets as showing that higher dystrophin levels are not associated with improved function in the eteplirsen trials.¹⁶

- Dr. Borio fundamentally agrees with Dr. Unger’s conclusion, and Dr. Jenkins concurs with Dr. Unger.¹⁷
- Dr. Woodcock finds that using “...the greatest flexibility possible for FDA while remaining within its statutory framework,” eteplirsen is “...reasonably likely to predict clinical benefit.”¹⁸ Her conclusion is based on a view that the data from both Study 201/202 and Study 301 are from adequate and well-controlled trials and that, although imperfect, they adequately meet criteria to include as affirmation of a drug effect that is reasonably likely to predict clinical benefit. She points out the many uncertainties about extrapolating from a particular level of a surrogate to clinical benefit when that surrogate is not yet proven, including complexities of assay validation, determining whether protein is functional, and also the extraordinary difficulty of knowing how the amount of protein might affect functional outcome over time and within the context of the multidimensional nature of protein interactions in complex cellular and subcellular functions. She finds no rational basis for identifying a specific threshold value for dystrophin levels that would be needed to support a determination that a particular level is “reasonably likely” to predict clinical benefit.¹⁹ Furthermore, she provides some post-hoc calculations from the eteplirsen clinical trials that she regards as supportive, though not definitive, evidence that higher levels of dystrophin are associated with greater function.

She is clearly employing and interpreting the full range of appropriate information, comprising a “totality of evidence” approach in determining that the clinical trials demonstrated an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit. Because of the uncertainties in this situation with a surrogate that has not been validated, it is clear that Dr. Woodcock’s decision also utilized the flexibility afforded under the relevant statutory provisions, including consideration of the life-threatening nature of the disease and the lack of alternative treatments.

Conclusions Regarding the Scientific Dispute

I conclude that qualified experts with extensive experience in FDA decision-making and stellar track records can assimilate the same scientific evidence and disagree about the extrapolation to whether the evidence supports a conclusion that the treatment has an effect that is “reasonably likely” to predict clinical benefit. Both Dr. Unger and Dr. Woodcock have drawn upon a deep level of knowledge and practical experience in deriving their conclusions, and their documents reflect strong convictions and profound concern both for the

¹⁶ *Id.* at 15-18.

¹⁷ SDR Board Recommendation at 15, footnote 94.

¹⁸ Woodcock Decisional Memorandum at 12.

¹⁹ *Id.* at 9.

well-being of the patients and families with DMD and for the preservation of the FDA's mission to protect and promote public health. That said, the history of the FDA includes a consistent precedent of final decision-making about medical products at the Center level. Overruling the Center Director is exceedingly rare and, in my view, would be appropriate only if the Center Director's decision could not be supported by the available data and information. In the present case, the scientific uncertainties lead to a situation in which the decision is a matter of reasoned expert opinion and judgment.

Given that I do not have technical expertise beyond those already involved in this decision and the record contains adequate evidence to support her conclusion, I defer to the judgment of the Center Director to approve eteplirsen under accelerated approval with the stipulations delineated in her Decisional Memo.

Additional Concerns

The SDR Board expresses concerns based on Dr. Unger's appeal and on interviews the Board conducted in its procedural review about the level of involvement of the Center Director in the review of the New Drug Application for eteplirsen. The following four points of concern were raised by Dr. Unger and cited by the SDR Board:

1. Intense involvement of the Center Director in early stages of review;
2. Extensive involvement in planning and participating in the Advisory Committee meeting to consider eteplirsen, held on April 25, 2016;
3. Initial decision by Dr. Woodcock on May 4, 2016 to approve before the review team had finalized its process for decision-making; and
4. Final decisional memorandum by Dr. Woodcock completed before Dr. Unger finalized his own decisional memo.

The SDR Board expresses concern about Dr. Woodcock's "extensive, early involvement in the review process" and states that "her involvement here appears to have upended the typical review and decision-making process."²⁰ Furthermore, the Board cautions that "...care should be taken to avoid the appearance of interfering with the integrity of scientific reviews at the lower levels of a Center."²¹ The SDR Board concludes that "Dr. Woodcock herself was the one who conducted that review and resolved the conflict in her own favor."²²

²⁰ SDR Board Recommendation at 23.

²¹ *Id.*

²² *Id.* at 27.

The concerns raised by Dr. Unger and by the SDR Board require that I evaluate the role of the Center Director. While there are many aspects of this situation that are unusual, leading to a rare formal dispute between a review team and a Center Director, I note the following points:

- Dr. Woodcock's management style has been "hands on" during her entire career at the FDA. She is well-known for interacting with staff at all levels and expressing her opinions. Indeed, complex and difficult cases often lead to early interactions with leaders, and as described below, frequent meetings and discussions along the way are common in CDER. This development program may represent a point of particular focus for Dr. Woodcock with the result that her involvement was more intensive than usual and correspondingly had an effect upon the review team, but such a focus does fit within a longstanding management approach.
- At the same time, under Dr. Woodcock's leadership and in conjunction with the leadership of Dr. John Jenkins in the Office of New Drugs, multiple mechanisms have been established for ongoing and interim discussions of evolving drug development programs and decision-making. These sessions, such as the Medical Policy Council, weekly senior staff meetings, the use of frequent Center Director briefings (in this case at least 14), and external advisory committee meetings are known for their exchange of ideas and welcoming of differences of opinion. Within this context it is also understandable that when so much is at stake with this degree of uncertainty, passionate debate can occur.
- An additional factor in this situation is the emergence of patient-centered drug development and the extensive interactions with the patient community as part of the overall environment for development and decision-making. While the appropriate methods for patient-centered drug development are evolving, the fact that DMD involves vulnerable children with a life-threatening illness and understandably concerned parents produces significant pressure on all involved. This dynamic is well reflected in Dr. Unger and Dr. Woodcock's documents. With a significant history dating back to the development of drugs for HIV/AIDS, patient-focused drug development is not an entirely new component of FDA's regulatory process, and it remains an explicit CDER priority in the current era.
- While others might direct and manage the organization and interact with patient and family communities in a different manner than Dr. Woodcock, it is difficult to argue that CDER has been unsuccessful under her leadership. A vast array of new drugs approved during her tenure is benefitting patients, including a host of significant advances using a variety of expedited programs. Further, this litany of therapeutic successes has been accomplished without compromising overall standards. Under Dr. Woodcock's leadership, commitment to due process and mutual review of scientific data have been major parts of CDER culture and are making a significant impact on drug development.

- It is inevitable that in some of these situations, highly qualified experts will disagree. In the face of profound changes in science and social interactions related to drugs, under Dr. Woodcock's leadership CDER enjoys broad public support, largely due to the constant, healthy atmosphere of transparency and debate that characterizes the documentation in this appeal. A review of Dr. Woodcock's record reveals extraordinary courage in the face of extreme pressure on many occasions, including from Congress, the press, patient and patient advocacy groups, and industry. She has taken and supported unpopular decisions when appropriate and is well-known for not relenting to pressure. Further, I find no pattern that indicates that this decision is part of a trend for lowering the standard for drug approval—a trend that would be unacceptable in any event.

Conclusions Regarding Additional Concerns

Overall, while I recognize the strain created by political and public pressures, given Dr. Woodcock's well-documented history of not bowing to such influences and a record in this case showing her close consideration of all relevant scientific evidence, **I do not find that she deviated from her responsibilities as Center Director, nor do I find that she succumbed to pressure from the patient community, the public, the press, or others.²³ Further, I do not find the general pattern of discourse and involvement to be atypical for Dr. Woodcock's management of the Center, nor do I find that her conduct was in conflict with the job requirements for Center Directors at the FDA.**

Additional Context: Accelerated Approval

It is important to note that the debate about this application is taking place in the context of an accelerated approval. Even under the best circumstances, accelerated approval may lead to decisions that are not verified upon further examination. These limitations are a reason accelerated approval is available only for a limited group of drugs: those intended to treat serious or life-threatening illnesses when the drug is expected to provide a meaningful benefit over existing therapy. In accelerated approvals, the surrogate is not a validated surrogate. As noted in FDA's Guidance on Expedited Programs for Serious Conditions – Drugs and Biologics, page 17:

For purposes of accelerated approval, a surrogate endpoint is a marker, such as a laboratory measurement, radiographic image, physical sign, or other measure, that is thought to predict clinical benefit, but is not itself a measure of clinical benefit. Depending on the strength of the evidence supporting the ability of a marker to predict clinical benefit, the marker may be a surrogate endpoint that is

²³ I was troubled by statements on page 16 of the SDR Board memo that Dr. Woodcock's decision to approve eteplirsen may have been inappropriately motivated by concerns over the sponsor's financial well-being. To address to my own satisfaction any questions that Dr. Woodcock's statements in the SDR Board memo might raise, I have discussed this issue directly with Dr. Woodcock, who said that she was aware of the financial pressures on the company, but that her decision was based on the science. Based on the record and our conversation, I am satisfied that her decision is indeed based on her scientific evaluation of the evidence.

known to predict clinical benefit (a validated surrogate endpoint that could be used for traditional approval), a surrogate endpoint that is reasonably likely to predict a drug's intended clinical benefit (and that could therefore be used as a basis for accelerated approval), or a marker for which there is insufficient evidence to support reliance on the marker as either kind of surrogate endpoint (and that therefore cannot be used to support traditional or accelerated approval of a marketing application).²⁴

Dr. Unger's and Dr. Borio's significant concerns about the implications of approving eteplirsen on the basis of the current data, even under the accelerated approval pathway, are evident.²⁵ As noted previously, Dr. Woodcock has explained that approval here will rely on "...the greatest flexibility possible for FDA."²⁶ The record amply reflects that the evidence here is not as strong as everyone involved in this dispute wishes it were. The record also reflects, however, that Dr. Woodcock has fully considered the patient population, lack of alternative therapies, and relevant information and analyses, and has concluded that the data are sufficient to meet the accelerated approval standard. Therefore, while I understand the concerns expressed by Drs. Unger and Borio regarding the implications of lowering of the standard for approval, I defer to Dr. Woodcock's conclusion that accelerated approval in this case does not represent such a lowering of the bar.

Opportunities for Process Improvement

I agree that there are many aspects of the history of eteplirsen that we must avoid in the future. Given the distinguished careers, time-proven expertise, and sincere concerns of Drs. Unger, Borio, Jenkins, and Woodcock, it is critical to review the following key lessons from this experience:

- All reviewers point to serious flaws in the eteplirsen drug development program. I personally have had many experiences in drug development and the assessment of therapeutics; it is common for problems to emerge that can be identified easily in

²⁴ More expansive discussion of these important issues on definitions of biomarkers and surrogates is included in a joint document from FDA and the National Institutes of Health, which can be found at: <http://www.ncbi.nlm.nih.gov/books/NBK326791>.

²⁵ In her capacity as Acting Chief Scientist, Dr. Borio notes: "Granting accelerated approval here on the basis of the data submitted could make matters worse for patients with no existing meaningful therapies—both by discouraging others from developing effective therapies for DMD and by encouraging other developers to seek approval for serious conditions before they have invested the time and research necessary to establish whether a product is likely to confer clinical benefit." Dr. Unger concludes: "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." Dr. Jenkins expressed the same concerns, both in written documents and in follow-up discussions I have had with him.

²⁶ Woodcock Decisional Memorandum at 12.

retrospect. Clinical development involves numerous human decisions and compromises engendered by the fact that clinical trials involve human research participants who must be respected independently of the many others who design, conduct and judge the research. Furthermore, the diverse interests of sponsors, patients and families, investigators, and regulators must be melded into a few protocols for human studies, so that no protocol represents 100% agreement. Accordingly, any retrospective criticism should be expressed with humility. However, the flaws in the eteplirsen program are serious and worthy of attention by others developing drugs for serious rare diseases.

- Specifically, the poor quality of many of the biopsies and the failure of the sponsor to implement a high-quality procedure for assay validation led to a situation in which only a fraction of the data could be used to make the regulatory decision. Given the serious nature of the disease and the invasiveness of the procedures, any studies must be conducted according to the highest standards, so that each child and parent who volunteers can be confident that they have contributed as much as possible to the generalizable knowledge needed to provide effective treatment based on high-quality evidence. Such attention to standards of quality in study conduct is a critical element of respect for research participants.
- Further, all reviewers agree that had the sponsor conducted properly-controlled, randomized trials with dose escalation from the beginning, we would by now likely have access to definitive information that would have resolved the disagreement. Dr. Woodcock points out that "... [t]here is no such thing as an 'exploratory study' for a serious, life-threatening disease without therapeutic options. Randomization should be performed very early in the development program and open-label studies should be avoided."²⁷ Time after time, we see the over-hyping of preclinical results playing into misguided arguments that prevent the generation of high-quality evidence because of unrealistic expectations of benefit, which ultimately fosters continued uncertainty about the clinical value of therapeutics.
- It is unfortunate that the sponsor touted an academically based study that had unreliable measures of the assay, thereby greatly overstating the degree of protein expression in the follow-up biopsies. Blinded experts assembled by the FDA fundamentally debunked this study, which has yet to be retracted and continues to be cited.²⁸ Dr. Woodcock identified the lack of a reliable assay as the primary defect in this "seriously deficient" development program, and Dr. Unger comments multiple times on the uncertainties in the assay and the manner in which poor methodology undermined the results presentation.²⁹ It is critical that we continue to improve the entire pipeline of translational research to introduce more rigorous and reliable methods, even in academically-conducted

²⁷ Woodcock Decisional Memorandum at 11.

²⁸ In view of the scientific deficiencies identified in this analysis, I believe it would be appropriate to initiate a dialogue that would lead to a formal correction or retraction (as appropriate) of the published report.

²⁹ *Id.*

translational research. Likewise, it is equally important that we continue to work on ways to purge discredited research in a manner that reduces inappropriate citation and follow-on research based on the incorrect assumption of the accuracy of the precedent studies.

- Dr. Unger also points out that data from primate studies suggest that a substantial increase in dose (approximately a log order) might well produce the 10% of normal dystrophin levels that Dr. Unger feels would meet criteria for an effect that is “reasonably likely to predict clinical benefit.”³⁰ This view that much higher doses are likely to be beneficial is shared across the entire FDA hierarchy, yet studies have not been done in humans at higher doses. However, the FDA must make a decision based on available data.
- As Drs. Unger, Woodcock, and Borio all point out, the risk of lowering the bar for new drug approval is a serious concern. We must remember that over 90% of drugs that enter clinical trials do not make it to market, usually due either to inadequate beneficial effect or unexpected toxicity.³¹ Lowering the regulatory hurdle could expose unsuspecting patients to the harmful effects of these drugs that would have failed if proper development had been done. But the statute and regulations clearly demarcate accelerated approval as a special situation in which a range of evidence from various sources and unvalidated surrogates may be used to determine whether the demonstrated effect on a surrogate endpoint is reasonably likely to predict clinical benefit, requiring an additional judgment on the part of the FDA. **Thus, I am confident that this unique situation will not set a general precedent for drug approvals under the accelerated approval pathway, as the statute and regulations are clear that each situation must be evaluated on its own merits based on the totality of data and information.**

The Path Forward

Given the approval status of eteplirsen, it is critical for the well-being of patients with DMD that the course of action fulfills the intent of accelerated approval and that the very best methods are used. As required for post-marketing studies for products approved under the accelerated approval program, the sponsor must conduct the required confirmatory trial with due diligence to evaluate whether eteplirsen has the predicted clinical benefit. It is notable that Dr. Unger and Dr. Woodcock agreed on several key points regarding further study of eteplirsen, regardless of the approval decision:

- The low doses currently employed are regarded as a starting point, as it seems likely that significant increases in dose could lead to desirable ranges of dystrophin production. Accordingly, a dose-finding study at much higher doses (as much as a log order increase)

³⁰ Appeal at 27.

³¹ Hay M, Thomas DW, Craighead JL, Economides C, Rosenthal J. Clinical development success rates for investigational drugs. *Nat Biotechnol.* 2014;32:40-51.

is needed to find the dose most likely to lead to maximal clinical improvement, as will be reflected in the approval letter for eteplirsen³²:

In order to verify the clinical benefit of eteplirsen, conduct a 2-year randomized, double-blind, controlled trial of eteplirsen in patients who have a confirmed mutation of the DMD gene that is amenable to exon 51 skipping. Patients should be randomized to the approved dosage of eteplirsen (30 mg/kg weekly) or to a dosage that provides significantly higher exposure, e.g., 30 mg/kg daily. The primary endpoint will be the North Star Ambulatory Assessment.

Dr. Unger points out that Study 301 could be converted to a randomized trial in which one arm remains on the current dosing regimen and the other(s) receive higher doses in the range that might be expected to produce much higher dystrophin levels.³³ A study with only a modest increase in dose that found no difference would be non-informative. Accordingly, the difference in doses in a post-approval trial should be enough to create clear separation in dystrophin levels.

- Dr. Woodcock notes that the sponsor is also conducting a randomized trial of eteplirsen versus no therapy in other forms of DMD.³⁴ The utmost attention should be paid to optimizing the methodological rigor of this trial and the trial evaluating higher doses discussed above.
- It is essential that the clinical and patient communities not be misled by exaggerated claims, given the poor quality of the studies in the sponsor's application and the fact that the approval is based on a non-validated surrogate and that clinical benefit has not yet been established. The applicable statutory and regulatory provisions governing accelerated approval, which require sponsors to submit promotional materials to FDA for pre-dissemination review, should be helpful in this regard.

Summary

In conclusion, I defer to Dr. Woodcock in her role as Center Director to make the decision to approve eteplirsen under the accelerated approval provisions. I find merit and reason on both sides of the disagreement, and despite the intensity of the argument, I believe that the quality of thinking on both sides reflects the importance of clinical and scientific expertise coupled with due process within the FDA. I find no basis for a view that Dr. Woodcock was unduly influenced by involvement with the patient community or other external pressures, and note that our understanding about how to include patients in the regulatory process is evolving. In addition, serious shortcomings present in the eteplirsen development program should not be allowed to establish a broad precedent for therapeutic development in rare diseases. In the end, we have

³² Draft Approval Letter for eteplirsen, as of September 9, 2016.

³³ Appeal at 27.

³⁴ Woodcock Decisional Memorandum at 10.

significant agreement on the science, but a disagreement on the integration of all the evidence as to whether the small changes in dystrophin levels induced by these doses of eteplirsen are “reasonably likely” to predict clinical benefit. Considering that a substantially flawed development program contributed to the difficulty coming to resolution in this case, we must redouble our efforts to move the therapeutic development ecosystem to use methods that will produce high-quality evidence from the outset.

For the reasons set forth above, CDER’s final decision, as set forth in Dr. Woodcock’s July 14, 2016 Memorandum, is upheld. This matter is remanded to CDER for action consistent with this decision.

 Sept. 16th, 2016

Robert M. Califf, M.D., Commissioner of Food and Drugs



Date: August 8, 2016

To: Robert Califf, M.D.
Commissioner of Food and Drugs

From: Luciana Borio, M.D.
Acting Chief Scientist

Luciana Borio -
A

Digitally signed by Luciana Borio -
DN: cn=US Government, ou=HHS,
ou=FDA, ou=People, cn=Luciana Borio -
A, o=FDA, ou=People, cn=Luciana Borio -
A, email=2016.08.08.20.00.37.0430*

Subject: Scientific Dispute Resolution Appeal regarding Eteplirsen

This matter is before the Office of the Commissioner on an appeal submitted by Ellis Unger, M.D., Director of the Office of Drug Evaluation I (ODE-I) (the initiator), under Staff Manual Guide 9010.1, “Scientific Dispute Resolution at FDA” (the SDR-SMG). In his scientific dispute resolution (SDR) appeal, dated July 18, 2016, Dr. Unger challenges the basis for a decisional memorandum issued by Janet Woodcock, M.D., Director of the Center for Drug Evaluation and Research (CDER). Dr. Woodcock’s decisional memorandum concludes that a new drug application (NDA) submitted by Sarepta Therapeutics Inc. (Sarepta) for eteplirsen, a drug intended to treat Duchenne muscular dystrophy (DMD), meets the standard for accelerated approval under 21 CFR § 314.510. Specifically, Dr. Woodcock’s memorandum states that the data submitted in support of the NDA establishes “increased dystrophin protein production, a surrogate endpoint [for DMD] that [she] conclude[s] is reasonably likely to predict clinical benefit.”¹ Dr. Unger states that he disagrees with Dr. Woodcock’s decisional memorandum because he does not believe “the findings on the dystrophin surrogate endpoint are reasonably likely to predict clinical benefit.”²

Upon receipt of the appeal from Dr. Unger, in accordance with the SDR-SMG, the Office of the Chief Scientist convened the Agency Scientific Dispute Process Review Board (the SDR Board), a standing committee, which I chair, whose role in evaluating the appeal is to conduct a review of the processes used in the Center to render a decision on the scientific dispute at issue.³ Under the SDR-SMG, “The goal of this review is to determine if the processes followed in the Center fully considered all relevant evidence and provided the initiator with an opportunity to express his or her concerns at all appropriate levels, prior to and including the Center Director.”⁴ My role in the process, as Chair of the SDR Board, is to provide a recommendation to you, as Commissioner of Food and Drugs, with respect to “whether a Center failed to follow its processes and/or did not provide an adequate opportunity to the initiator to express his or concerns; [whether] all relevant evidence bearing on the scientific question at issue has been considered; and[] whether the dispute should be remanded to the Center Director.”⁵ The written

¹ Woodcock Decisional Memorandum at 1.

² Appeal at 3.

³ SDR-SMG at 3. (“The Agency Scientific Dispute Process Review Board (hereafter Board) is a standing committee comprised of representatives of the Office of Accountability and Integrity, Ombudsmen from all Centers and the agency (or officials so designated) and representative(s) from the Office of the Chief Scientist. The Board is chaired by the Chief Scientist.”).

⁴ *Id.* at 12.

⁵ *Id.* at 5.

recommendation must reflect the SDR Board’s underlying rationale, along with minority views among the members, for those findings.⁶

In conducting its evaluation, the SDR Board reviewed pertinent aspects of the Center’s administrative file for the eteplirsen NDA and interviewed Dr. Unger, Dr. Woodcock, one member of the review team for the NDA, who requested anonymity, and Virginia Behr, the Ombudsman for CDER. Based on its review, the SDR Board has determined that the processes followed by CDER provided Dr. Unger with an adequate opportunity to present his scientific views and that CDER considered all relevant evidence. As Chair of the SDR Board, I therefore recommend that you do not remand this matter to the Center Director for further action.⁷ However, there are additional considerations meriting your attention, which I describe below. Furthermore, the SDR Board encourages you to conduct a thorough substantive review of the scientific dispute in this matter or, in the alternative, to convene a panel of relevant experts to conduct such a review and provide advice to the agency and you, as Commissioner, on whether the evidence of the effect of eteplirsen on the surrogate endpoint is reasonably likely to predict clinical benefit.

BACKGROUND

1. *Eteplirsen and DMD*

Dr. Unger provides an overview of eteplirsen and DMD in his appeal.⁸ In short, DMD is a genetic disorder with catastrophic effects on its sufferers:

[DMD] is an X-linked recessive neuromuscular disorder caused by mutations of the dystrophin gene[,] . . . [which] disrupt the messenger ribonucleic acid (mRNA) reading frame [and] lead[] to the absence or near-absence of dystrophin protein in muscle cells. . . . Absence of dystrophin leads to muscle damage, with replacement by fat and collagen. . . [and a concomitant] loss of physical function in childhood and adolescence, with premature death from respiratory and/or cardiac failure in the second to fourth decade.⁹

There are no FDA-approved therapies for DMD.¹⁰ Sarepta has designed eteplirsen to target the pre-mRNA transcripts of the dystrophin gene so that exon 51 is excluded from the resulting mRNA:¹¹

[B]y restoring [] the mRNA reading frame, a ‘truncated’ but nevertheless partially functional form of the dystrophin protein can be produced by muscle cells, delaying disease progression. Similar truncated dystrophin is found in a less severe form of muscular dystrophy, Becker Muscular Dystrophy (BMD). In essence, the drug is hoped to induce production of sufficient Becker-type dystrophin to slow the progression of the disease. This drug is specific for exon 51 mutations, a subset of the mutations that cause DMD. If approved, the drug

⁶ *Id.* at 13.

⁷ *See id.* (“The Commissioner will review the [SDR Board’s] recommendation and render a final decision on . . . whether the dispute should be remanded to the Center Director for corrective action” and “work with the Center Director to determine what corrective actions must be taken, if any.”).

⁸ Unless otherwise indicated, Drs. Unger and Woodcock appear to agree as to the background provided in this section.

⁹ Appeal at 2.

¹⁰ *Id.*

¹¹ The charity, Muscular Dystrophy UK, has a nice description of the technology underpinning eteplirsen, which can be accessed at: <http://www.muscular dystrophyuk.org/progress-in-research/background-information/what-is-exon-skipping-and-how-does-it-work/>.

would be indicated for ~13% of the overall DMD patient population. Eteplirsen has not received marketing authorization from any regulatory authority, and no similar drugs are approved.¹²

In attempting to establish that eteplirsen is safe and effective for the treatment of DMD, and thus meets one of the standards for approval in the Federal Food, Drug, and Cosmetic Act (FD&C Act), Sarepta has submitted data from three clinical studies:

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

Dr. Unger further explains:

The endpoints for [the three] studies can be broadly divided into those that aim to show changes in physical performance, e.g., walking speed, rise time from the floor, muscle function; and those that aim to show effects on production of dystrophin in skeletal muscle – the surrogate endpoint. Dystrophin was quantified in this development program using two methods: Western blot and immunohistochemistry.¹⁴

Immunohistochemistry (IHC) analysis looks at thin slices of muscle biopsies to see if dystrophin is present or absent. Each muscle fiber that shows any amount of dystrophin is counted as positive, regardless of the actual quantity of dystrophin present. Western blot analysis assesses how much dystrophin is present.

For Study 201/202, Sarepta submitted Western blot and IHC analysis evaluating proteins in muscle samples obtained from the twelve patients before the study and then again at twelve, 24, and 48 weeks.¹⁵ “The Western blots submitted by the applicant for Study 201 were oversaturated, unreliable, and uninterpretable.”¹⁶ Because CDER also determined that the conditions under which the original IHC analysis was performed were inadequate, including that the reader was not masked to sequence and time, the Center requested a re-reading of the stored images by three masked pathologists under different conditions.¹⁷ The IHC results from the reread were not nearly as favorable, as compared to the initial IHC results reported by Sarepta.

¹² Appeal at 2.

¹³ *Id.*

¹⁴ *Id.*

¹⁵ *Id.* at 4, 8.

¹⁶ *Id.* at 4.

¹⁷ Unger Decisional Memorandum at 12-13.

The re-read showed a nominally statistically significant increase in dystrophin in response to eteplirsen for the low dose group, but not the high dose group[] (. . . [T]he type-I error rate was not controlled for multiplicity.)¹⁸ Moreover, for the 4 patients who had received placebo through Week 24 and then switched to eteplirsen, there was no increase in dystrophin at Week 48.¹⁹

For Study 201/202, CDER also worked with Sarepta to improve the Western blot assays, and researchers performed repeat biopsies on eleven of twelve patients at Week 180.²⁰ Only three of the eleven patients had stored baseline samples that were adequate for evaluation, and so baseline samples were obtained from six additional patients external to Study 201/202.²¹ Dr. Unger also notes that all baseline samples were obtained from a different muscle group than the samples obtained at Week 180.²² Based on its own analysis of the IHC data, Sarepta claimed a remarkable increase of dystrophin immunostaining at Week 180: from 1.1% \pm 1.3% positive muscle fibers at baseline to 17.4% \pm 10.0% positive fibers at Week 180.²³ The Western blot analysis resulted in Week 180 dystrophin levels that were small, with a mean increase of only 0.93% of normal dystrophin levels in the muscle fibers.²⁴ Dr. Unger remarked that the lack of concordance between the IHC and the Western Blot results is “striking” and also noted that FDA did not verify the integrity of the IHC results.²⁵ As previously noted, each muscle fiber that shows any amount of dystrophin is counted as positive in IHC, regardless of the actual quantity of dystrophin present.

As noted above, Study 301 is an ongoing study. For purposes of its review of the NDA, CDER requested that Sarepta perform Western blot analysis on samples obtained from 13 patients enrolled in the study.²⁶ The analysis compared paired biceps samples: baseline samples and samples obtained at 48 weeks, after 48 weeks of treatment with 30 mg/kg of eteplirsen infusion.²⁷ Dr. Woodcock told the SDR Board that representatives from CDER were present in the laboratory for the Western blot analysis and oversaw the procedures and controls. The Western blot analysis showed a statistically significant increase in dystrophin, ranging in an increase from 0.22% to 0.32% of normal.²⁸ It should be noted, however, that a statistically significant increase in dystrophin, the surrogate endpoint, of an exceptionally small magnitude does not imply clinical benefit, which is the issue at the core of Drs. Unger and Woodcock’s scientific disagreement.

¹⁸ That is, with respect to time points of assessment and the 2 doses tested.

¹⁹ Appeal at 8-9. Of note, in her decisional memorandum, Dr. Woodcock rejected the findings in both the original and second evaluation of the images: “Much of the controversy over the adequacy of these assessments relates to the fact that rigorously validated assays were not used to evaluate the initial 3 muscle biopsies, apparently resulting in overestimation of the various readouts and some irreproducibility of IHC and Western blot dystrophin assays. For these reasons, I do not discuss or rely upon the results of these earlier assays, or on re-reads of them.” (Woodcock Decisional Memorandum at 2). She explained to the SDR Board that, after consultation with others in CDER, she does not view IHC results standing alone as a valid method to evaluate dystrophin levels.

²⁰ Appeal at 5.

²¹ *Id.* at 5, 9.

²² *Id.* at 5. Dr. Unger clarifies in his decisional memorandum that the baseline biopsies were from the biceps muscle, the Week 180 biopsies from the deltoid muscle. (Unger Decisional Memorandum at 17).

²³ Appeal at 9. As discussed below, however, Dr. Unger does not believe that those results are reliable.

²⁴ *Id.* at 5.

²⁵ *Id.* at 9-10.

²⁶ *Id.* at 6. Dr. Unger states that the biopsies were obtained from 13 patients but only reports the data as to 12 patients. “There was one patient for whom none of the values met the acceptance criteria [for the Western blot assay].” (Unger Decisional Memorandum at 21).

²⁷ Appeal at 6.

²⁸ *Id.*

2. *Legal Standard for Accelerated Approval and Patient Perspectives*

On December 11, 1992, on the basis of its broad statutory authority to approve drugs under the FD&C Act, FDA issued regulations providing for accelerated approval of drugs.²⁹ Under 21 CFR § 314.510, FDA may grant accelerated approval for a drug based on a surrogate endpoint under certain circumstances:

FDA may grant marketing approval for a new drug product on the basis of adequate and well-controlled clinical trials establishing that the drug product has an effect on **a surrogate endpoint that is reasonably likely**, based on epidemiologic, therapeutic, pathophysiologic, or other evidence, **to predict clinical benefit** or on the basis of an effect on a clinical endpoint other than survival or irreversible morbidity. Approval under this section will be subject to the requirement that the applicant study the drug further, to verify and describe its clinical benefit, where there is uncertainty as to the relation of the surrogate endpoint to clinical benefit, or of the observed clinical benefit to ultimate outcome. Postmarketing studies would usually be studies already underway. When required to be conducted, such studies must also be adequate and well-controlled. The applicant shall carry out any such studies with due diligence.³⁰

The preamble to the proposed rule defines “surrogate endpoint” as follows:

A surrogate endpoint, or “marker,” is a laboratory measurement or physical sign that is used in therapeutic trials as a substitute for a clinically meaningful endpoint that is a direct measure of how a patient feels, functions, or survives and that is expected to predict the effect of the therapy. For example, elevated cholesterol and hypertension, two surrogate endpoints, are important because they are risk factors for coronary and cerebral artery disease; but it is the impact of the diseases (e.g., angina, congestive heart failure after a heart attack, paralysis after a stroke, or sudden death) that is important to the patient.³¹

In 2012, Congress passed the Food and Drug Administration Safety and Innovation Act (FDASIA). Section 901 of FDASIA amended the FD&C Act to provide FDA with specific authority to grant accelerated approval to drugs for serious conditions.³² Section 506(c) of the FD&C Act now largely tracks language in the regulations issued by FDA in 1992. Section 901 of FDASIA also added current section 506(e) to the FD&C Act, which clarifies that the amendments were “intended to encourage [FDA] to utilize innovative and flexible approaches to the assessment of products under accelerated approval” but that “[n]othing 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 [the FD&C Act].”³³

Section 901 of FDASIA also directed FDA to issue guidance to industry on the development of

²⁹ 57 Fed. Reg. 58942 (Dec. 11, 1992).

³⁰ Emphasis added.

³¹ 57 Fed. Reg. 13234, 13235 (Apr. 15, 1992).

³² FDASIA, PL 112-144, July 9, 2012, 126 Stat. 993.

³³ *Id.*

drugs for accelerated approval and required consideration of the following:

In developing the guidance . . . [FDA] shall consider how to incorporate novel approaches to the review of surrogate endpoints based on pathophysiologic and pharmacologic evidence in such guidance, especially in instances where the low prevalence of a disease renders the existence or collection of other types of data unlikely or impractical.³⁴

Section 1137 of FDASIA further directs FDA to:

develop and implement strategies to solicit the views of patients during the medical product development process and consider the perspectives of patients during regulatory discussions, including by—(1) fostering participation of a patient representative who may serve as a special government employee in appropriate agency meetings with medical product sponsors and investigators; and (2) exploring means to provide for identification of patient representatives who do not have any, or have minimal, financial interests in the medical products industry.³⁵

In May 2014, FDA finalized a guidance on “Expedited Programs for Serious Conditions — Drugs and Biologics.” The Guidance provides general information on the evidence that the agency considers in determining whether to grant accelerated approval.³⁶ The Guidance clarifies that assessing a surrogate endpoint hinges on understanding both the disease process and the relationship between the drug’s effect and the disease process.³⁷ With respect to the latter, the Guidance states:

The extent to which a drug’s effect on the surrogate endpoint is known to predict an effect on the disease either because the effect is on the causal pathway or correlates with clinical outcomes is critical. Sometimes this relationship can be assessed epidemiologically[,] but it is most persuasively established by knowing that a drug that affects the surrogate endpoint also affects a clinical outcome.³⁸

The Guidance also provides some insight on how the agency exercises its judgment in evaluating surrogate endpoints when little is known about how an effect on a surrogate endpoint might affect clinical endpoints:

Particularly in rare diseases, there may be limited information in the literature, lack of in-depth epidemiological or historical data, and little or no experience with other drugs to inform the interpretation of surrogate endpoints or intermediate clinical endpoints. FDA may consult with external experts on surrogate endpoints and intermediate clinical endpoints where there is a lack of historical data for a given disease.³⁹

³⁴ *Id.*

³⁵ *Id.*

³⁶ Expedited Programs Guidance at 19-22.

³⁷ *Id.* at 20-22.

³⁸ *Id.* at 21.

³⁹ *Id.* at 21-22.

FDA obtains patient perspectives through a variety of avenues, “such as open public hearings on specific diseases or drug development issues, and as speakers at FDA-sponsored conferences and workshops.”⁴⁰

3. *SDR-SMG and CDER’s SDR-SOPs*

The Office of the Commissioner issued the SDR-SMG on January 13, 2009. Its stated purpose is “to improve the process of internal scientific dispute resolution[] and to encourage open communication throughout the agency.”⁴¹ The SMG “encourages the resolution of scientific disputes at the working level in the organization, starting with the frontline employees and their immediate supervisors or team leaders” and cautions that the “agency’s appeals process for scientific disputes is not a replacement for robust and fair Center-level processes.”⁴² As noted above, the SDR-SMG provides for submission of SDR appeals to the Office of the Commissioner and outlines the process and standards for evaluating such appeals. Under the SDR-SMG, the SDR Board evaluates whether “the processes followed in the Center fully considered all relevant evidence and provided the initiator with an opportunity to express his or her concerns at all appropriate levels, prior to and including the Center Director.”⁴³ As Chair of the SDR Board, the Chief Scientist then provides a written recommendation on those issues to the Commissioner, who renders a final decision on whether the scientific dispute should be remanded to the Center for further action.⁴⁴

In addition to outlining the process for elevating scientific disputes to the Office of the Commissioner, the SDR-SMG details the agency’s “requirements for the minimum standards for scientific dispute resolution processes in the Centers” and provides a collection of non-mandatory “best practice[s]” for such dispute resolution.⁴⁵ The SDR-SMG’s requirements for resolving scientific disputes at the Center-level begin with an obligation on the part of Center management to ensure open scientific debate on controversial issues:

Center management shall create an atmosphere in which consultation and open discussion on controversial issues are encouraged. When disagreements occur, it is necessary to follow appropriate procedures for resolving them. Informal methods, using good management practices for resolving conflict, should be employed prior to instituting the more formal procedures described here. Notwithstanding informal good management practices used to try to resolve the conflict, timely written reviews of the scientific matter in dispute should be completed by all members of a review group, including initiator and supervisors, to enable as open and complete a discussion of the issues as possible at the working level of the organization.⁴⁶

The SDR-SMG then goes on to require the Centers to have in place written standard operating procedures for formally resolving scientific disputes (SDR-SOPs) in the event that such informal attempts at resolution are unsuccessful.⁴⁷ In contrast to the procedural review contemplated by

⁴⁰ 79 Fed. Reg. 65410, 65411 (Nov. 4, 2014).

⁴¹ SDR-SMG at 1.

⁴² *Id.* at 2.

⁴³ *Id.* at 12.

⁴⁴ *Id.* at 12-13.

⁴⁵ *Id.* at 2-3.

⁴⁶ *Id.* at 6.

⁴⁷ *Id.*

the SDR-SMG, Center-level SDR-SOPs should provide for substantive review of the scientific disputes at issue within the Center.⁴⁸

At CDER, there are three interrelated chapters of the Center’s Manual of Policies and Procedures (MAPPs) that serve to implement the SDR-SMG’s requirements. The first, MAPP 4151.8, “Equal Voice: Discipline and Organizational Component Collaboration in Scientific and/or Regulatory Decisions,” sets forth CDER’s principles for resolving scientific disputes informally and requires “a collaborative environment for decision-making.”⁴⁹ According to the MAPP, “[s]uch an environment requires open communication and exchange of ideas in a mutually respectful professional environment[] and the full and open participation of all relevant disciplines and organizational components in the decision-making process.”⁵⁰ MAPP 4151.8 states that “[e]ach individual who contributes to the decision-making process” must “be sure the position represented is consistent with the scientific, regulatory, and/or administrative policies of that . . . organizational component” and that “[o]pinions of staff should be documented and supported by data in a matter commensurate with the magnitude of the decision being made.”⁵¹

The second and third MAPPs at issue directly relate to CDER’s formal SDR process. MAPP 4151.1, “Scientific/Regulatory Dispute Resolution for Individuals Within a Management Chain,” provides for raising a scientific issue to the “Next Highest Management Official” (NHMO) if alignment on an issue cannot be reached by the staff on a team or through discussions with a team leader or first-level supervisor. The individual who disagrees with the decision (the disputant) “. . . may initiate a dispute resolution process by writing a statement (called a dispute statement) describing the position, concept, opinion, or recommendations with which the disputant disagrees . . . as well as the proposed changes and rationale for the changes in recommendations and/or conclusions.”⁵²

The disputant submits the statement to the NHMO, i.e., “the management official one level above the management official who made the decision being disputed.”⁵³ The NHMO then issues a written decision on the issue, and any disputant may then appeal the written decision up the chain of command all the way to the Center Director through use of the same process.⁵⁴ MAPP 4151.2, “Resolution of Differing Professional Opinions: Review by Ad Hoc Panel and CDER Director,” provides for further formal review under certain circumstances if alignment cannot be reached under the process in MAPP 4151.1.⁵⁵ A CDER employee may initiate the process by submitting a written package, which must include “[a]n assessment of the possible significant negative consequences to the public health” at issue in the dispute, to the CDER Ombudsman.⁵⁶ The CDER Ombudsman and the Center Director then “determine whether the consequences of the decision in question are potentially serious enough to warrant” additional review.⁵⁷ If so, the Center Director appoints a chairperson to lead an *Ad Hoc* review panel for purposes of evaluating the scientific dispute and providing a recommendation to the Center

⁴⁸ See *id.*; see also footnote 136.

⁴⁹ MAPP 4151.8 at 2.

⁵⁰ *Id.*

⁵¹ *Id.* at 2-3.

⁵² MAPP 4151.1 at 3.

⁵³ *Id.*

⁵⁴ *Id.* at 4.

⁵⁵ *Id.* at 5; MAPP 4151.2 at 1-2.

⁵⁶ MAPP 4151.2 at 5.

⁵⁷ *Id.*; see also *id.* (“In most cases, the Ombudsman will ensure that all other avenues for resolution (e.g., dispute resolution process, Advisory Committee discussion, CDER regulatory briefing) have been exhausted . . .”).

Director, who renders the final decision.⁵⁸ The *Ad Hoc* panel typically includes one member with relevant technical expertise, one member chosen from a list provided by the person requesting review, and, if possible, one member with relevant expertise who is external to the agency.⁵⁹

4. *Procedural History of the Dispute in CDER*

Sarepta submitted its NDA for eteplirsen (#206488) on June 26, 2015.⁶⁰ CDER assigned it for review to the Division of Neurology Products (DNP) within ODE-I, the office for which Dr. Unger serves as Director.⁶¹ Even before submission of the NDA, however, representatives from the Office of New Drugs (OND), DNP and ODE-I (the review team) regularly briefed Dr. Woodcock on issues related to the ongoing study of eteplirsen pursuant to an investigational new drug application (IND) and the anticipated NDA.⁶² The discussions at these briefings included among their topics: the suitability of eteplirsen for accelerated approval, an overview and background for eteplirsen, study design, a clinical site inspection report for Sarepta, general brainstorming, and planned communications.⁶³ Dr. Unger told the SDR Board both that there were far more briefings of the Center Director than is typical and that the scope of those briefings included an unusual level of detailed discussion.

During the SDR Board’s separate interviews of Dr. Unger and the review team member (RTM), the SDR Board learned that, at Dr. Woodcock’s direction, the review team also joined her in meetings with patient advocacy groups for DMD on multiple occasions—anywhere from six to twelve times—from very early on in the review process. The RTM described the meetings with the patient advocacy groups, which frequently included boys with DMD and their parents, as “intense,” “personal,” and “intimidating.” Dr. Unger and the RTM both thought that Dr. Woodcock’s early interest and involvement in DNP’s approach to guiding the development of eteplirsen was based in part on the enthusiasm in the DMD community in relation to an article published about the initial findings for Study 201/202, which Drs. Unger and Woodcock now agree are misleading and unreliable. Indeed, Dr. Woodcock told the SDR Board that she became involved because of the broader public interest the article generated, along with encouragement from the Commissioner of Food and Drugs at the time and her long-held belief that OND has been very conservative in evaluating drugs for accelerated approval. In his decisional memorandum, Dr. Unger explains the excitement surrounding eteplirsen at the time as follows:

[The initial findings for Study 201/202] were substantially reported in a 2013 publication, 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≤0.002). Even greater increases occurred at week 48 (52% and 43%*

⁵⁸ *Id.* at 6-7.

⁵⁹ *Id.* at 6.

⁶⁰ Unger Decisional Memorandum at 1.

⁶¹ *Id.* at 2.

⁶² Appeal at 24-25; Chronology prepared by Virginia Behr and submitted to the SDR Board (Behr Chronology) at 1-2. In his appeal, Dr. Unger consistently refers to the representatives from OND, OND-I and DNP who were involved in the review of the eteplirsen NDA as the “review team” or as “the division,” even though he appears to be referring to senior management within OND on occasion. Dr. Woodcock has also used the same terminology on occasion, though not as consistently. For the sake of efficiency, this memorandum refers to everyone at CDER who was involved in the review of the eteplirsen NDA, besides Dr. Woodcock herself, as the review team. Nonetheless, the SDR Board notes that, within FDA, “review team” is often used to reflect the core team of individuals within a division who are directly engaged in the review of the science underlying a regulatory submission.

⁶³ Appeal at 24-25; Behr Chronology at 1-2.

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, which also claimed a remarkable treatment effect from eteplirsen and raised wildly unrealistic expectations in the DMD community.⁶⁴

In their interviews with the SDR Board, Dr. Unger and Dr. Woodcock stated that FDA also received significant correspondence from the public and Congress, much of which urged approval of eteplirsen.⁶⁵ Some of the correspondence used vulgar language and was abusive to the review staff.⁶⁶

The briefings of Dr. Woodcock began again five to six months after submission of the NDA for eteplirsen.⁶⁷ The focus of these briefings was on preparation for a planned meeting of the Peripheral and Central Nervous System Drugs Advisory Committee (AC meeting) to provide advice on the review of the eteplirsen NDA, which meeting was initially scheduled for January 2016 but then rescheduled for April 25, 2016.⁶⁸ The preparation involved discussions of the ongoing review of the data, including the “strengths, limitations, and uncertainties of the data, particularly with respect to the comparison between the open-label eteplirsen group and a contemporary untreated external control group.”⁶⁹ During their respective interviews with the SDR Board, both Dr. Unger and the RTM conveyed their belief that Dr. Woodcock was inclined to grant approval from very early on in the process. But the RTM stated that Dr. Woodcock’s views were not always clear during discussions throughout the review of the science—sometimes she seemed to agree with external constituents, sometimes not. The RTM told the SDR Board that, in his or her view, the review team was never sure whether they were discussing science, policies, or politics. According to both Dr. Unger and the RTM, Dr. Woodcock frequently conveyed that she thought the review team was being unreasonable and encouraged DNP to find a way to approve the eteplirsen NDA. Both Dr. Unger and the RTM told the SDR Board that Dr. Woodcock seemed focused on the external pressures, from both patient advocacy groups and Congress, and that she frequently talked about the effects of a decision regarding eteplirsen in terms of overarching policy (e.g., the need to be more flexible for ultra-rare diseases). The RTM highlighted to the SDR Board that at least two members of the review team were leaving FDA or had left the agency in the wake of both the decision-making process within CDER and the pressures exerted by outside forces.

Dr. Woodcock conceded to the SDR Board that she was leaning toward granting approval in light of the available data as early as 2014. She said that her goal throughout the discussions

⁶⁴ Unger Decisional Memorandum at 11 (emphasis in original), citing Mendell JR, *et al*: Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* 2013;74:637-47 and Sarepta press release, dated 8/8/13 (<http://investorrelations.sarepta.com/phoenix.zhtml?c=64231&p=irolnewsArticle&ID=1846052>). Dr. Unger also notes, “It was these perceptions and expectations that led the applicant to declare that a placebo-controlled study was no longer feasible.” Unger Decisional Memorandum at 11.

⁶⁵ See also Appeal at 23.

⁶⁶ See, e.g., *id.* at 23-24.

⁶⁷ *Id.* at 25; Behr Chronology at 2-3.

⁶⁸ Appeal at 25; Behr Chronology at 2-3.

⁶⁹ Appeal at 25.

with the review team was to convince them to come around to her more flexible way of thinking about the data. According to Dr. Woodcock, she recognized that there were serious and significant flaws in the study design for Study 201/202 and the data it generated but that she did not “want to hold” those flaws “against the patients.” She conceded that the results produced by Studies 201/202 and 301 were always less than anyone in CDER had hoped.

In their respective interviews with the SDR Board, both Dr. Unger and the RTM focused to some extent on Dr. Woodcock’s involvement in the planning stages for the AC meeting. They expressed some surprise at the extent of her involvement. Dr. Unger indicated in his interview with the SDR Board that Dr. Woodcock even advocated, unsuccessfully, for changing the order of the questions to be posed to the committee and wanted the question on conventional approval to come before the one on accelerated approval.

The RTM told the SDR Board: (1) that Dr. Woodcock made it clear in one or more of the meetings leading up to the AC meeting that she intended to speak at the meeting but (2) that the substance and purpose of her participation were never communicated. Although the RTM affirmatively stated that the review team was free to develop its own presentation to the committee, uncertainty with respect to Dr. Woodcock’s role made doing so more difficult. The RTM also noted that Dr. Woodcock requested a longer than is typical Open Public Hearing portion of the AC meeting that, as a result, the review team thought there would insufficient time for them to make their presentations during a one-day meeting. The RTM stated that the review team asked to extend the advisory committee to two days but that they were overruled.

On April 25, 2016, CDER held the AC meeting. The meeting focused on the data from Study 201/202.⁷⁰ Dr. Woodcock spoke at the meeting several times. At the meeting she made a presentation that was intended to “provide a framework within which to consider [the] data [underlying the eteplirsen NDA] based on [her] 30 years of experience at FDA and really extensive experience in implementation of the legal standards for drug approval.”⁷¹ She highlighted many of the difficulties in interpreting the data.⁷²

At the AC meeting, Dr. Woodcock also described the standards for both conventional and accelerated approval of drugs but mentioned that the agency had not “articulated an evidentiary standard for determining if a surrogate endpoint is reasonably likely to predict clinical benefit.”⁷³ She concluded her presentation with the following remarks:

I would note that much of the effort in evaluating a drug development program goes into avoiding a specific mistake, that is erroneously approving a drug that is not effective.

There often is little consideration of another error, which is failing to approve a drug that actually works. In devastating diseases, the consequences of this mistake can be extreme, but most of these consequences are borne by patients who traditionally [] have little say in how the standards are implemented.

The accelerated approval program includes a requirement for confirmatory studies for efficacy, so as you've heard from the sponsor, you have to do further studies to explore and confirm effectiveness. An inherent

⁷⁰ Sarepta had not yet submitted the data from Study 301.

⁷¹ Advisory Committee Transcript at 151.

⁷² *Id.* at 151-155.

⁷³ *Id.* at 155-156.

presumption in this program of accelerated approval, which is written in the preamble to our regulation about it, is that more uncertainty is going to be tolerated initially and that in fact sometimes we will collectively get it wrong, otherwise accelerated approval would really have no different standards than regular approval.⁷⁴

During the questions to the committee members, Dr. Woodcock restated the standard for accelerated approval and emphasized that, with regard to the surrogate endpoint of dystrophin, there has never been a “threshold established [to show a reasonable likelihood of predicting clinical benefit] because there's never been a drug to do this.”⁷⁵ When later asked for clarification of the extent to which the committee members were to incorporate the testimony of the boys and their families into their evaluation of clinical outcomes for Study 201/202, Dr. Woodcock stated:

Well, we are instructed, as people said, to take the use of the patient community into account, more on the benefit and the risk. * * * So the statutory standard is more or less as described there, but there is flexibility, and that's where we should take the views of the community into account.⁷⁶

During his SDR Board interview, the RTM stated that, notwithstanding Dr. Woodcock's emphasis on accelerated approval and the standard of “reasonably likely to predict clinical benefit,” “[s]urrogacy was not discussed in any genuine scientific way” during the AC meeting because it had not been framed that way by Sarepta through its presentation to the committee. The RTM specifically stated that there was no discussion of “substantial evidence” in the context of accelerated approval, nor what might constitute “interpretable evidence.” The RTM believed that, by the end of an emotional AC meeting, the framework for evaluating the data under the appropriate regulatory standards, as provided by the review team toward the start of the meeting, had been forgotten by the committee members.

Dr. Woodcock explained to the SDR Board that she thought both that the review team did a poor job framing the issues during their presentations and that the questions were confusing and poorly worded. Indeed, during her interview with the SDR Board, Dr. Woodcock opined that the review team “did not put its best foot forward.” She speculated that the confounding factor was the number of interested persons attending both in person and by webcast. She stated that she did not interfere with either aspect of the AC meeting because she knew she disagreed with the review team and Dr. Unger had already signaled that he would file an SDR appeal if she decided to grant accelerated approval to eteplirsen. She thought that the review team's presentation of the IHC data, in particular, was confusing. She further opined that the review team's failure to highlight the clinical data made the questions on conventional approval and accelerated approval difficult for the committee members to understand. Dr. Woodcock also criticized the review team for how it downplayed and undercut the views of the patient advocates.

At the conclusion of the AC meeting, the committee voted against accelerated approval by a margin of 7-6.⁷⁷ Three of the members who voted in favor of accelerated approval were the consumer representative and the two patient representatives.⁷⁸

⁷⁴ *Id.* at 158-59.

⁷⁵ *Id.* at 484.

⁷⁶ *Id.* at 548-549.

⁷⁷ *Id.* at 486-95.

⁷⁸ *Id.* at 2-7, 486-88.

On May 4, 2016, Dr. Woodcock met with the review team to discuss the AC meeting and plan of actions for the NDA.⁷⁹ In his appeal, Dr. Unger contends that Dr. Woodcock “made clear her intent to approve the drug” at this meeting, even though she had not yet reviewed drafts of DNP’s final review memorandum or his review memorandum.⁸⁰ According to Dr. Unger, Dr. Woodcock explained that she had already “reached a different conclusion” than the review team.⁸¹ Dr. Woodcock explained to the SDR Board that the memoranda were discussed during the Center Director briefings and that she felt she understood the views of the review team and did not see the point of an “exchange of reviews.”

On May 24, 2016, Dr. Unger met privately with Dr. Woodcock to discuss the eteplirsen decision.⁸² On May 31, 2016, Dr. Woodcock met with representatives from the review team to discuss their reviews and her initial draft of a decisional memorandum based primarily on the data from Study 201/202.⁸³ Dr. Woodcock received comments back from the review team at the same meeting.⁸⁴ Dr. Unger told the SDR Board that he and members of the review team—including Dr. Robert Temple, Deputy Center Director for Clinical Science and Dr. John Jenkins, Director of OND—discouraged Dr. Woodcock from finalizing the decisional memorandum and granting accelerated approval for eteplirsen until the additional data from Study 301 could be obtained.

On June 3, 2016, in response to an email from Sarepta, a letter signed by Dr. Woodcock issued to the sponsor.⁸⁵ The letter requested the additional data from Study 301, which was to include comparisons of any biopsy samples obtained at Week 48 to the respective baseline samples for those patients.⁸⁶ The letter stated,

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 within four business days of receiving the data (assuming all other aspects of the application are approvable).⁸⁷

Dr. Woodcock explained that Dr. Unger and the review team essentially agreed to the timeframe of four business days, though they pushed instead for six. She felt that there was general agreement that data from only twelve patients could be reviewed quickly, especially given that representatives from CDER would be overseeing the Western blot analysis and ensuring that it was done properly.

On June 27, 2016, Sarepta submitted the requested data.⁸⁸ Dr. Woodcock explained that accelerated approval was not granted within four business days of that date precisely because the results of the analysis were disappointing in that they provided evidence of only a minimal increase in dystrophin at 48 weeks. Dr. Unger sent an email to Dr. Woodcock that read:

⁷⁹ Appeal at 25; Behr Chronology at 2.

⁸⁰ Appeal at 26.

⁸¹ *Id.*

⁸² Behr Chronology at 2.

⁸³ Appeal at 26; Behr Chronology at 2.

⁸⁴ Behr Chronology at 2.

⁸⁵ June 3, 2016, General Advice letter.

⁸⁶ *Id.* at 1.

⁸⁷ *Id.* at 1-2.

⁸⁸ Unger email to the SDR Board, dated July 22, 2016.

I don't have to tell you how difficult the eteplirsen decision has been for many of us in ODE-I. As you know, we have reached different scientific conclusions on the strength of the data, and in particular, the likelihood that the small increase observed in Becker-type dystrophin is reasonably likely to predict clinical benefit. This decision could be precedent setting with respect to accelerated approval, i.e., where the bar should be set for changes in a pharmacodynamic biomarker that are deemed "reasonably likely to predict clinical benefit." Moreover, to my knowledge, this could be the first time a Center Director has overruled a review team (and an advisory committee) on a question of whether effectiveness has been demonstrated.

I know that Dr. Jenkins has mentioned the possibility of involving Dr. Califf in the eteplirsen decision on at least one occasion, and I would like to request a formal appeal to the Commissioner on this matter.

I'm aware that the Commissioner's official role is to consider the administrative aspects of review decisions and not the science. But given the potential for setting a precedent here, I think he should be aware of the various points of view and consider the potential ramifications of the matter at hand.

I'm also aware that you advised Sarepta that we would be prepared to grant accelerated approval of their NDA within 4 business days of receiving their new data, but there was a provision in the letter that the increase in dystrophin had to be meaningful, and we do not have agreement on this point. Thus, it is my hope that a Commissioner Briefing can be held before an action is taken.

I have discussed the above with Dr. Jenkins, and he supports this course of action.

I propose that we reserve a few minutes at the briefing tomorrow to discuss this matter.⁸⁹

On July 6, 2016, Dr. Woodcock met with the review team one final time.⁹⁰ During the meeting, Dr. Woodcock "indicated to the review team that [she] had read their memoranda that had been updated to reflect the new [Western blot] data, and that [she] maintained [her] position that the application should receive accelerated approval based on dystrophin production."⁹¹ She discussed her rationale, which—based on her notes—appears to have tracked the rationale in her final decisional memorandum.⁹²

On July 8, 2016, in light of Dr. Unger's stated intention of filing an appeal with the Office of the Commissioner, Virginia Behr, CDER Ombudsman, began working with him and Dr. Woodcock to determine whether the institution of any formal appeals under CDER's SDR-SOPs was warranted.⁹³ Ms. Behr had determined that the procedure outlined in MAPP 4151.1,

⁸⁹ Unger email dated July 5, 2016.

⁹⁰ Appeal at 26; Behr Chronology at 3.

⁹¹ Woodcock's handwritten notes, dated July 6, 2016, at 1.

⁹² *Id.* at 2. Also of note, on July 7, 2016, Dr. Unger briefed you on his rationale for disagreeing with Dr. Woodcock's underlying scientific reasoning for granting accelerated approval for eteplirsen (Behr Chronology at 3).

⁹³ See "Agreement to utilize FDA Staff Manual Guide 9010.1 for internal appeal related to NDA 206488, eteplirsen injection" (SDR-SOPs Agreement).

“Scientific/Regulatory Dispute Resolution for Individuals Within a Management Chain” did not apply because the disagreement was between the Center Director and a subordinate two levels below her.⁹⁴ She also questioned the utility of using MAPP 4151.2, “Resolution of Differing Professional Opinions: Review by Ad Hoc Panel and CDER Director.”⁹⁵ She reasoned that “the CDER Director ha[d] already fully evaluated the issues and [was] one of the parties involved in the dispute” and that “utilizing this MAPP could potentially extend this already lengthy NDA action another 50 business days.”⁹⁶ She nonetheless consulted with both Drs. Unger and Woodcock, who both agreed to bypass the *Ad Hoc* panel process in favor of the process outlined in the SDR-SMG.⁹⁷ During his presentation to the SDR Board, Dr. Unger also indicated that he thinks referring the matter to an *Ad Hoc* panel would have been pointless because Dr. Woodcock had already made up her mind and a new process would not have changed the outcome.

On July 11, 2016, Dr. Woodcock provided a draft of her final decisional memorandum to the review team.⁹⁸ She received comments back from Dr. Unger; Dr. Jenkins, the Director of OND; and Dr. Ashutosh Rao, of the Office of Biotechnology Products, who was also on the review team.⁹⁹ The comments from Drs. Unger and Rao do not debate the action proposed in Dr. Woodcock’s draft decisional memorandum or its underlying scientific conclusions.¹⁰⁰ Instead, they focus on clarifying certain facts asserted in the memorandum, and Dr. Unger provided information regarding the clinical course of 11 patients enrolled in Study 201/202 to 240 weeks.¹⁰¹ Dr. Jenkins provided more detailed analysis on and critique of some of Dr. Woodcock’s findings and he expressed concern about her conclusions. However, he made no attempt in his written comments to dissuade her from her ultimate conclusion regarding accelerated approval.¹⁰² By email on the afternoon of July 13, Dr. Unger stated, “I’ve canvassed the Division, and we have no additional comments.”¹⁰³ Dr. Unger told the SDR Board that he and the review team understood that Dr. Woodcock had already made up her mind and that thus they did not see a point in criticizing Dr. Woodcock’s draft decisional memorandum.

Furthermore, the RTM told the SDR Board that some of the positions taken by Dr. Woodcock in the draft decisional memorandum were brand new to him but that he did not feel any feedback he could provide would receive due consideration by Dr. Woodcock. The RTM expressed concern that Dr. Woodcock’s analysis for “reasonably likely to predict clinical benefit” raised new issues and information that should have been presented at the beginning of the review and that had not been addressed by the review team or, perhaps more importantly, presented by the sponsor in support of the NDA. The RTM specifically discussed with the SDR Board the section of the finalized version of the memorandum addressing whether the data for eteplirsen is adequate to show a reasonable likelihood of predicting clinical benefit.¹⁰⁴ As an example of his concerns, the RTM pointed to section (B)(5) of the decisional memorandum, which details the findings in the

⁹⁴ *Id.* at 1. It is also clear from the record before the SDR Board that the supervisor between Drs. Unger and Woodcock, Dr. John Jenkins, agreed with Dr. Unger.

⁹⁵ *Id.*

⁹⁶ *Id.* at 2.

⁹⁷ *Id.*

⁹⁸ Behr Chronology at 3.

⁹⁹ Unger emails dated July 13, 2016 and sent at 12:57 AM, 9:26 AM (including attachments), and 11:19 AM (including attachment); Jenkins email dated July 12, 2016; and emails (including attachments) from Rao dated July 12 and 13, 2016.

¹⁰⁰ Unger emails dated July 13, 2016 and sent at 12:57 AM, 9:26 AM (including attachments), and 11:19 AM (including attachment); emails (including attachments) from Rao dated July 12 and 13, 2016.

¹⁰¹ Unger emails dated July 13, 2016 and sent at 12:57 AM, 9:26 AM (including attachments), and 11:19 AM (including attachment); emails (including attachments) from Rao dated July 12 and 13, 2016.

¹⁰² Jenkins email dated July 12, 2016.

¹⁰³ Unger email dated July 13, 2016 and sent at 3:19 PM.

¹⁰⁴ Woodcock Decisional Memorandum at 5-10.

scientific literature regarding “the relationship of dystrophin expression to clinical status.”¹⁰⁵

The RTM indicated that he or she knows the scientific literature at issue very well and that he or she could have provided significant input into the evaluation of the literature and the underlying data and analysis. The RTM conveyed that he did not do so because he felt Dr. Woodcock had already made her decision.

On July 14, 2016, Dr. Woodcock finalized her decisional memorandum. She explained to the SDR Board that her conclusion regarding whether the increase in dystrophin production identified by Studies 202 and 301 was reasonably likely to predict clinical benefit was based on her own “medical/scientific judgment.” She emphasized that she has thirty years of experience at FDA and that she has far more experience in assessing this type of evidence for an “ultra-rare rare” disease than the review team. She thought that the review team was unreasonable in its position on a threshold for predicting clinical benefit in this case. Her stated goal for the decisional process was to move the review team toward what she viewed as a more reasonable approach. She acknowledged that there were clear weaknesses in the data but that accelerated approval should not be limited to “sure bet” drugs and that confirmatory trials are required for a reason. Dr. Woodcock emphasized her view that the agency needs to accept more uncertainty when granting accelerated approval. She also criticized OND for not issuing clear guidance on what constitutes a sufficient drug effect to be “reasonably likely to predict clinical benefit,” as she had suggested for an extended period of time. She also thought that the review team’s views on balancing the mean results of a clinical study with a targeted evaluation of responsive patients were misplaced, particularly in a DMD population, where additional genetic mutations or deficiencies could have a profound effect on the outcome.

In her presentation to the SDR Board, Dr. Woodcock suggested that, in making the decision, she was looking at the broader picture for the development of these types of drugs for very limited patient populations in the United States (between 600 and 1300) and that there needed to be some path forward for such innovative products. She opined that Sarepta in particular “needed to be capitalized.” She noted that the sponsor’s stock went down after the AC meeting and went up after FDA sent the June 3, 2016 letter. Dr. Woodcock cautioned that, if Sarepta did not receive accelerated approval for eteplirsen, it would have insufficient funding to continue to study eteplirsen and the other similar drugs in its pipeline. She stated that, without an approval in cases such as eteplirsen, patients would abandon all hope of approval for these types of products and would “lapse into a position of” self-treatment.

On July 16, 2016, Dr. Unger finalized his own decisional memorandum. In her own decisional memorandum, dated July 14, 2016, Dr. Woodcock indicated that she had read Dr. Unger’s decisional memorandum,¹⁰⁶ although she could not have done so given the timing of the two memoranda. She explained to the SDR Board that she did not feel she needed to see a finalized version of Dr. Unger’s decisional memorandum because she was already familiar with his views on the data and the decision. She also stated that there was nothing in Dr. Unger’s appeal, which is based largely on his finalized decisional memorandum, that would have changed her mind on her decision or the underlying rationale. She stated, “He is entitled to his own opinion.”

5. Dr. Unger’s SDR Appeal

In his appeal, Dr. Unger focuses his arguments almost exclusively on the substance of his scientific disagreement with Dr. Woodcock. Indeed, Dr. Unger makes clear in his appeal that he

¹⁰⁵ *Id.* at 7-10.

¹⁰⁶ *Id.* at 1.

seeks “a scientific review on the matter of whether or not there is substantial evidence of a quantitative effect on dystrophin protein that is reasonably likely to predict clinical benefit.”¹⁰⁷ Insofar as he explicitly addresses potential procedural issues under the review process contemplated by the SDR-SMG, he does so in two paragraphs toward the end of the appeal.¹⁰⁸ He first states that Dr. Woodcock’s “direct involvement with this drug, compared to other development programs, has been unprecedented.”¹⁰⁹ He states further that “[s]he also attended the April meeting of the Peripheral and Central Nervous System Drugs Advisory Committee, where she spoke and interjected a number of important comments.”¹¹⁰ After conceding that “[t]here is no question that there has been adequate time and place for the discussion of various views,” Dr. Unger notes that he found it unfortunate that “the Center Director made clear her intent to approve the drug at a briefing with the review team on May 4, 2016, before she had seen drafts of the Division’s final review memorandum or my review memorandum.”¹¹¹ As noted above, Dr. Unger indicates that Dr. Woodcock conveyed that she had “already ‘ . . . reached a different conclusion . . . ’ than the review team.”¹¹²

In his presentation to the SDR Board, Dr. Unger highlighted that Dr. Woodcock had never seen the charts on page 10 of his appeal. Those charts show: (1) a comparison of the *original* IHC results for baseline samples in the three patients whose biopsies were available at 180 weeks to the IHC results for those same samples when they were re-evaluated after 180 weeks and (2) a comparison of the IHC and the Western blot results at 180 weeks.¹¹³ Dr. Unger stated, however, that those charts were consistent with his earlier positions and would likely not affect Dr. Woodcock’s analysis or decision. In a follow-up email to the SDR Board, Dr. Unger also contended that Dr. Woodcock diverted from protocol when she finalized her decisional memorandum on July 14, 2016, two days before his.

In his appeal, Dr. Unger frames his scientific disagreement with Dr. Woodcock as follows: “The disagreement is over the question of whether the findings on the dystrophin surrogate endpoint are reasonably likely to predict clinical benefit.”¹¹⁴ Nonetheless, Dr. Unger explains his disagreement with Dr. Woodcock through multiple challenges to the reliability of the underlying data and specific issues he has with her rationale or the evidentiary basis for such rationale. Of note, he makes the following scientific arguments:

- As noted above, Study 201 showed only “a nominally statistically significant increase in dystrophin in response to eteplirsen for the low dose group...”¹¹⁵
- Study 201/202 was fundamentally flawed in several respects:
 - “[T]he baseline biopsies were obtained from [external controls] . . . who could differ in unknown ways from the subjects in Study 201/202”;¹¹⁶
 - “[T]he Week 180 biopsies were obtained from different muscles than the baseline biopsies”;¹¹⁷ and

¹⁰⁷ Appeal at 26.

¹⁰⁸ *Id.*

¹⁰⁹ *Id.*

¹¹⁰ *Id.*

¹¹¹ *Id.*

¹¹² *Id.*

¹¹³ *Id.* at 10.

¹¹⁴ *Id.* at 3.

¹¹⁵ *Id.* at 9.

¹¹⁶ *Id.* at 5.

¹¹⁷ *Id.*

- “The baseline biopsies for the three subjects with Week 180 data had been stored for several years and the protein may have degraded, leading to a falsely low baseline value, and a greater apparent increase from baseline....”¹¹⁸
- Although the available data generated by Study 301 were the product of an adequate and well-controlled study and showed a statistically significant increase of dystrophin, the drug effect (i.e., an increase from 0.22% to 0.32% of normal) is not reasonably likely to predict clinical benefit.¹¹⁹
 - “The treatment effect observed cannot be compared or related to levels of dystrophin measured by other laboratories and reported in various publications”;¹²⁰
 - “Dr. Woodcock never provides a rational argument – based on reliable data – to support the concept that ‘...low-level increases in dystrophin production are reasonably likely to predict clinical benefit.’ She provides no rationale – no link between a mean increase in dystrophin of 3 parts per thousand and clinical benefit”;¹²¹ and
 - “No evidence of a clinical effect was demonstrated in the eteplirsen development program, and there is no correlation between dystrophin levels as determined by Western blot and clinical outcome.”¹²²

He also makes several overarching policy and legal arguments that call into question the appropriateness of Dr. Woodcock’s decisional memorandum. His key arguments focus on the effects that Dr. Woodcock’s decision would have on the pathway for accelerated approval and the standard for “reasonably likely to predict clinical benefit.”¹²³ He also highlights the negative effects that accelerated approval would have on the patients themselves, including false hope, abandonment of other therapies, and a decline in drug development for DMD.¹²⁴ He further questions “the ethics of approving or prescribing a drug for a fatal disease at a dose that is very likely to be sub-therapeutic[] when the consequence of a sub-therapeutic dose is clinical deterioration and death.”¹²⁵ Finally he worries that approving eteplirsen based on the data submitted by the sponsor “would send the signal that political pressure and even intimidation—not science—guide[] FDA decisions.”¹²⁶

ANALYSIS

1. *Whether CDER followed its own processes.*

The first issue for the SDR Board to consider is whether CDER followed its own processes in addressing Dr. Unger’s scientific dispute. Dr. Unger does not contend that there were any issues with respect to how CDER chose to address and implement its own formal appeals process under the SDR-SOPs in this case. In his appeal, Dr. Unger points instead to four deviations from typical Center process: (1) Dr. Woodcock’s involvement in the early stages of review of the eteplirsen NDA; (2) her extensive involvement in planning the AC meeting and her participation

¹¹⁸ *Id.*

¹¹⁹ *Id.* at 7.

¹²⁰ *Id.* at 13.

¹²¹ *Id.* at 15.

¹²² *Id.*

¹²³ *Id.* at 21-22.

¹²⁴ *Id.*

¹²⁵ *Id.* at 23; *see also* Unger Review Memorandum at 4, 5.

¹²⁶ *Id.*

in the meeting; (3) her initial decision (on May 4, 2016) to approve the eteplirsen NDA before the review team had completed even their draft review memoranda; and (4) her issuance of her final decisional memorandum before Dr. Unger finalized his own decisional memorandum as Director of ODE-I. In its review of the administrative file and the surrounding circumstances, the SDR Board has also identified below other potential deviations from process at the Center level.

The agency-wide SDR-SMG directs the SDR Board to focus on the Center’s SDR-SOPs in evaluating whether the Center followed its own processes in evaluating a scientific dispute. In this case, however, both Drs. Unger and Woodcock have agreed that the only applicable SDR-SOP, MAPP 4151.2, provides for a review by the Center Director in consultation with an *Ad Hoc* panel and that going through such a process at this stage would be futile. The SDR Board has determined that, absent the second aspect of that agreement regarding futility and the underlying unusual circumstance of this scientific dispute, there would be reason to refer the matter back to the Center for further review by an *Ad Hoc* panel.

The interplay between MAPP 4151.1 and 4151.2, the former of which provides for supervisory review of scientific disputes all the way to the Center Director, suggests that MAPP 4151.2 actually calls for additional review of a scientific dispute by the Center Director under certain circumstances even if she has already made a decision on the dispute. Although MAPP 4151.2 provides for bypassing review of the scientific dispute up the chain of command under MAPP 4151.1 if such exhaustion would impede the timely resolution of a serious public health issue, MAPP 4151.2 also emphasizes that it should not be used before other means of resolution have been attempted.¹²⁷ However, the key consideration for obtaining review by an *Ad Hoc* panel under MAPP 4151.2 is “whether the consequences of the decision in question are potentially serious enough to warrant [additional review],” not whether the resort to the process would be futile.¹²⁸ It appears that Dr. Woodcock has never made a determination regarding the seriousness of the decision in question, but it would be surprising if she determined that the dispute in this case did not meet the standard, as reflected in the statement she signed.¹²⁹

In this case, however, it is clear from the record before the SDR Board that Dr. Woodcock was so involved in the underlying scientific dispute—including direct and extensive personal review of the data and analyses offered in support of the NDA—that we agree with the conclusion in the agreements signed by Drs. Unger and Woodcock that “the CDER Director has already fully evaluated the issues.”¹³⁰ Indeed, she has already received advice from an advisory committee and had substantial conversations with her staff over an extended period of time with respect to the dispute in question. There is no reason to believe that receiving additional advice from an *Ad Hoc* panel would alter Dr. Woodcock’s views of the scientific issues. As the agreement between her and Dr. Unger reflects, the process would be time-consuming and delay an important

¹²⁷ MAPP 4151.2 at 5. (“In most cases, the Ombudsman will ensure that all other avenues for resolution (e.g., dispute resolution process, Advisory Committee discussion, CDER regulatory briefing) have been exhausted before a [request for review under 4151.2] is filed. However, in some cases, an individual may believe that his or her professional opinion will not be considered by his or her supervisors or that there is not time to exhaust other options for dispute resolution without seriously endangering the public health. In this case, the submitter should include . . . a written request to bypass these other mechanisms. . . .”).

¹²⁸ *Id.*

¹²⁹ SDR-SOPs Agreement at 2 (“The difference of opinion between Drs. Unger and Woodcock could be considered to meet the criteria for filing an appeal under MAPP 4151.2 because the drug indication sought is one for a serious and life-threatening disease that has limited treatment options.”).

¹³⁰ *Id.*

regulatory decision unnecessarily.¹³¹ Dr. Unger also told the SDR Board that he thought going through the *Ad Hoc* panel process would have been pointless for the aforementioned reasons.

The difficulty for the SDR Board is that the agency-wide SDR-SMG is predicated on some level of formal scientific dispute resolution within the Center, particularly a decision by the Center Director regarding the formalized scientific dispute.¹³² For that reason, the focus of the SDR-SMG with respect to the process followed is on whether the Center followed its own SDR-SOPs in resolving the scientific dispute.¹³³ Yet, the SDR-SMG also directs the Centers to adopt “[i]nformal methods” for resolving scientific disputes, “to create an atmosphere in which consultation and open discussion on controversial issues are encouraged,” to use “good management practices for resolving conflict,” and “to enable as open and complete a discussion of the issues as possible at the working level of the organization.”¹³⁴ As a result, the SDR Board has determined that reviewing the processes used by a Center to resolve a scientific disagreement is appropriate under the SDR-SMG even when, as here, the initiator has not availed himself of the Center’s formal process for resolving scientific disputes and the Center Director has explicitly agreed to that approach.

Whether the Center followed its own processes for resolving a scientific disagreement cannot be viewed in a vacuum, however. Indeed, the SDR-SMG itself—at its most concise and in its clearest voice—states, “The goal of [the SDR Board’s] review is to determine if the processes followed in the Center fully considered all relevant evidence and provided the initiator with an opportunity to express his or her concerns at all appropriate levels, prior to and including the Center Director.”¹³⁵ Particularly in the context of a scientific dispute that did not go through a formal SDR process at the Center but nonetheless received extensive review by the Center Director, focusing on deviations from process without any regard to whether they affected the initiator’s opportunity to present his views of the science (and to some extent whether those views and the evidence were considered) would seem to miss the point of that review. Accordingly, the SDR Board finds that it is more appropriate to address Dr. Unger’s arguments regarding the Center’s deviations from appropriate process under the second prong of its analysis: whether the Center provided Dr. Unger an adequate opportunity present his scientific concerns.

The SDR Board’s one caveat is that, as noted above, the SDR-SMG does appear to assume that there has been both at least some use of the formal dispute resolution within the Center and, accordingly, a *formal* substantive review of the initiator’s scientific concerns before reaching the Office of the Commissioner.¹³⁶ The limited scope of the SDR Board’s review under the SDR-SMG—i.e., an evaluation of the Center’s decision-making process—means that Dr. Unger will also not receive a substantive review of his scientific concerns under the SDR-SMG. In fact, at the conclusion of the SDR Board’s review, Dr. Unger will not have received a substantive review of his scientific concerns under any formal process at any level. Particularly in light of

¹³¹ *Id.* (“[U]tilizing this MAPP could potentially extend this already lengthy NDA action another 50 business days.”).

¹³² *See* SDR-SMG at 6 (requiring as a mandatory process for formal scientific dispute resolution a written opinion by the Center Director and stating that such a written opinion as a step in the process is a “central criterion for advancement to the agency-level appeals process.”).

¹³³ *See, e.g., id.* at 12 (requiring the SDR Board to “obtain the full administrative record of the Center’s processes for the dispute and review the Center’s published SOP(s)” and to “review that information to determine whether written Center processes were followed.”).

¹³⁴ *Id.* at 6.

¹³⁵ *Id.* at 12.

¹³⁶ *See id.* at 6 (referring to SOPs for resolution of Center-level scientific disputes without limiting them to procedural reviews and contemplating the Center SOPs as a continuation of the informal SDR process).

Dr. Unger's explicit request for scientific review of the matter within the Office of the Commissioner, therefore, the SDR Board recommends additional substantive review at this level, as is discussed below.

2. *Whether CDER provided an adequate opportunity to Dr. Unger to present his scientific concerns.*

In his appeal, Dr. Unger admits, "There is no question that there has been adequate time and place for the discussion of various views."¹³⁷ In so doing, he appears to concede away most of his arguments with respect to whether he had an adequate opportunity to present his scientific concerns, notwithstanding the procedural deviations he identifies. The SDR Board, however, has not taken Dr. Unger's concession at face value and has instead looked beyond it to evaluate the administrative file and the surrounding circumstances to identify additional procedural issues. We conclude nonetheless that Dr. Unger had an adequate opportunity to present his scientific concerns to Dr. Woodcock before she issued her decisional memorandum.

As noted above, Dr. Unger identified four deviations from Center's typical decision-making process for the eteplirsen NDA: (1) Dr. Woodcock's involvement in the early stages of review of the eteplirsen NDA; (2) her extensive involvement in planning the AC meeting and her participation in the meeting; (3) her initial decision (on May 4, 2016) to approve the eteplirsen NDA before the review team had completed even their draft review memoranda; and (4) her issuance of her final decisional memorandum before Dr. Unger finalized his own decisional memorandum as Director of ODE-I. In reviewing this matter, the SDR Board—which includes among its members Ombudsmen from other Centers that oversee reviews of medical products—also considered other departures from the typical processes used by Centers in reviewing applications for pre-market approval or clearance.¹³⁸

The SDR Board agrees with Dr. Unger that it was unusual for a Center Director to be so involved in the early stages of reviewing an NDA, but the consensus on the SDR Board was that Dr. Woodcock went several steps further than mere involvement and thereby departed from typical practice among the Centers. By her own admission, Dr. Woodcock had a direct hand in reviewing the data submitted in support of the NDA, even before the review team had written their draft review memoranda, and actively encouraged the review team—including Dr. Unger—to come around to her way of thinking in their own reviews. Specifically, she wanted the review team to agree with her that the limited increase in dystrophin production established by the data in Studies 201/202 was sufficient to show a reasonable likelihood of predicting clinical benefit. At several points during the decision-making process for what is clearly a critical scientific issue for the agency, Dr. Woodcock also provided a very limited amount of time for Dr. Unger and the review team to provide feedback on additional data or her own scientific conclusions—most notably when Sarepta submitted the data from Study 301 and when she provided two separate draft versions of her decisional memorandum to the review team.

Notwithstanding the foregoing procedural shortcomings, the SDR Board finds that Dr. Unger had an adequate opportunity to present his scientific views. Not only does he admit in his appeal that he had an opportunity, but the record before the SDR Board demonstrates that he did. He and the rest of the review team met with Dr. Woodcock on multiple occasions both before and after the AC meeting. Drs. Unger and Woodcock both told the SDR Board that those meetings involved substantive and detailed discussions of the data and science and the appropriate

¹³⁷ Appeal at 26.

¹³⁸ See SDR-SMG at 3 (defining the SDR Board to include Ombudsmen from all of the Centers).

conclusions to be drawn from them. Although Dr. Unger complains that Dr. Woodcock was involved in aspects of the NDA that went far beyond the norm for a Center Director at CDER, including her role in the AC meeting, and that she reached or finalized decisions before reviewing review or decisional memoranda, he does not maintain that those procedural deficiencies compromised his ability to present his views. In fact, his own final decisional memorandum—which Dr. Woodcock apparently saw in draft form before she finalized her own—discloses that he felt empowered to push back on both Dr. Woodcock’s scientific conclusions and their basis, despite the fact that he believed his efforts would be futile. Indeed, he conceded to the SDR Board that nothing in his decisional memorandum or appeal submission would have affected Dr. Woodcock’s decision on the scientific issue in question (including the charts that he created for the first time in preparing his appeal submission under the agency-wide SDR-SMG). He further conceded as much when he agreed not to pursue further review through the *Ad Hoc* panel process under CDER’s SDR-SOPs. In short, through his own perseverance, confidence in his own scientific expertise, and perhaps dint of personality, Dr. Unger ensured that he himself had an adequate opportunity to present his scientific views despite the procedural irregularities in the decision-making process within CDER.

The SDR Board nonetheless remains concerned about Dr. Woodcock’s extensive involvement in the review of the eteplirsen NDA, including her degree of participation at the AC meeting, and the limited timeframe she provided for feedback on the data from Study 301 and her own scientific conclusions on that data. We fear that those actions could have chilled scientific debate within CDER and reduced the level of participation by the review team during the final stages of the decision-making process. By all accounts, Dr. Woodcock made clear her views that CDER should lean toward finding that eteplirsen met the standards underlying accelerated approval nearly from the outset of her involvement. By May 4, 2016, she had orally communicated her intention to grant accelerated approval for eteplirsen, even though she had not yet seen even the draft review memoranda from the review team or a decisional memorandum from Dr. Unger. Then, when she requested data from Study 301 from Sarepta, she communicated to the sponsor a compressed timeframe for CDER’s review. Although she later expanded the timeframe for review when the data proved to be disappointing, she apparently analyzed the data on her own, conducted her own additional search of the scientific literature, and took only six or seven business days to orally communicate to the review team her decision to grant approval.

To complicate matters further, Dr. Woodcock subsequently circulated a draft decisional memorandum but provided only a limited amount of time for comments, even though the draft decisional memorandum was the first time some on the review team had apparently seen key elements for the basis of her decision on “reasonably likely to predict clinical benefit.” The response from the review team is telling. As noted above, only Drs. Jenkins and Unger and another reviewer outside of DNP provided comments. Except for Dr. Jenkins, no one made any effort to make substantive comments beyond tips on how to make factual clarifications or to supplement her analysis with additional data. It appears that, because the review team knew Dr. Woodcock’s views by then, they saw no point in providing any additional substantive review or meaningful feedback on any new issues raised by Dr. Woodcock’s memorandum. Indeed, Dr. Unger and the RTM conveyed as much to the SDR Board.

There is no doubt that a Center Director should have wide latitude in leading the direction of the Center in a manner consistent with her priorities and vision. The SDR Board also believes that Center Directors have a role to play not only with respect to the resolution of scientific disputes at issue in individual applications for pre-market-authorization by FDA, as evidenced by both the SDR-SMG and CDER’s own SDR-SOPs, but also with respect to the ultimate decision on

scientific issues that are not the subject of a dispute. It is also clear from Dr. Woodcock’s presentation to the SDR Board that she firmly believes in the correctness of her scientific decision in this case and that her involvement in the review of the eteplirsen NDA was always motivated by the best of intentions. However, the SDR Board finds Dr. Woodcock’s extensive, early involvement in the review process troubling. Indeed, her involvement here appears to have upended the typical review and decision-making process.

Rather than ensuring that the scientific reviews started at the bottom of the chain of command, Dr. Woodcock made clear from her position at the top that she was pushing for a particular outcome from the very early stages. As a consequence, the regulatory reviews did not start at the staff level with scientific reviews and then proceed through the chain of command for concurrence or non-concurrence at all appropriate levels within the management structure, as would be the typical course of decision-making for a regulatory decision grounded in science. Indeed, before the reviewers had even completed their draft scientific reviews, Dr. Woodcock had told them—on May 4, 2016—that she intended to grant accelerated approval. This sort of top-down review does not, in the SDR Board’s view, “create an atmosphere in which consultation and open discussion on controversial issues are encouraged,” as reflected in the SDR-SMG’s requirements for resolution of scientific disagreements by the Center.¹³⁹ By the time Dr. Woodcock issued her draft decisional memorandum on what she herself acknowledged was a difficult scientific issue of incredible magnitude for the agency—i.e., whether the evidence regarding dystrophin production was reasonably likely to predict clinical benefit—the review team had decided it was pointless to challenge her ultimate conclusion or its basis.¹⁴⁰ Review teams should have the opportunity to conduct their reviews without preemption by the Center Director. As noted above, the SDR Board believes that Center Directors should have a role in shaping policy, expressing concerns, and resolving issues once they are ripe for their review, but we caution that care should be taken to avoid the appearance of interfering with the integrity of scientific reviews at the lower levels of a Center.

3. Whether the Center Director considered all relevant evidence bearing on the scientific question at issue.

The third issue for the SDR Board is whether CDER, including Dr. Woodcock, fully considered all relevant evidence in resolving the scientific dispute at issue, i.e., whether the evidence of eteplirsen’s effect on dystrophin production is reasonably likely to predict clinical benefit. In this case, both Drs. Unger and Woodcock appear to agree that she *considered* all relevant evidence. As noted above, Dr. Unger does not believe that any additional data or evidence available to him could persuade Dr. Woodcock that she has reached the wrong scientific conclusion. For her part, Dr. Woodcock does not feel that she has disregarded any relevant evidence. Moreover, in her interview with the SDR Board, she demonstrated an awareness and command of all of the evidence weighing against the scientific decision she has made, including the arguments and analysis of the evidence presented in Dr. Unger’s appeal.

Whether Dr. Woodcock has *addressed* all of the relevant evidence in her decisional memorandum is a more difficult question. In concluding that the minimal increase in dystrophin

¹³⁹ *Id.* at 6.

¹⁴⁰ In this regard, it is also worth noting again the language quoted above in the background section: “Each individual who contributes to the decision-making process” must “be sure the position represented is consistent with the scientific, regulatory, and/or administrative policies of that . . . organizational component” and that “[o]pinions of staff should be documented and supported by data in a matter commensurate with the magnitude of the decision being made.” (MAPP 4151.8 at 2-3).

production seen in the data is reasonably likely to predict clinical benefit, Dr. Woodcock has provided a very limited rationale.

At the risk of oversimplification, Dr. Woodcock found, in essence, that the studies attempting to correlate levels of dystrophin with clinical benefit, as have been reported in the scientific literature, are unreliable in this context for variety of reasons, including: (1) the subjectivity of the clinical evaluation, (2) the difficulty in correlating IHC results with Western blot results, (3) the influence of anti-dystrophin antibodies, (4) the lack of information on dystrophin quality (as opposed to quantity) in the different studies, (5) deficiencies in Western blot techniques from earlier studies, and (6) the wide range of findings with respect to the correlation of dystrophin levels with clinical benefit.¹⁴¹ She concluded, therefore, that “protein in the range between undetectable and 10% of normal is likely to be very important for clinical presentation, all other things being equal, i.e. mutation status and non-dystrophin-related factors affecting phenotype,” and that the “biochemical data strongly support the idea that low-level increases in dystrophin production are reasonably likely to predict clinical benefit.”¹⁴² She then attempted to bolster that conclusion with a theory regarding the effect of exon 52 deletion and her reanalysis of the intermediate clinical outcomes for a subset of subjects in Study 201/202.¹⁴³ She further explained to the SDR Board that she was exercising her “medical/scientific judgment” in reaching the scientific conclusion that she did.

It is easy for the SDR Board to understand why Dr. Unger’s appeal expressed such frustration with this explanation of Dr. Woodcock’s rationale. He states:

I believe the burden is on Dr. Woodcock to show or explain why production of a near-zero quantity of dystrophin (0.3%) is reasonably likely to predict clinical benefit, and I do not believe her July 14, 2016 memo makes this case. I believe that the available evidence leaves open the possibility that some patients could benefit from a small increase in dystrophin, but this possibility does not reach the threshold of being reasonably likely to predict a clinical benefit.¹⁴⁴

Of course, *considering* the relevant evidence and *addressing* the relevant evidence in a manner satisfactory to Dr. Unger or the SDR Board are two different propositions. The SDR Board finds, based on the record before us, that Dr. Woodcock has considered all relevant evidence in reaching her scientific conclusion. Based on her own medical judgment, she simply has a difference of opinion with Dr. Unger—both with respect to the scientific conclusion and the sufficiency of the underlying rationale.

4. Whether the dispute should be remanded to the Center Director.

Inasmuch as the SDR Board has concluded that Dr. Unger had an adequate opportunity to present his scientific concerns during the decision-making process at CDER and that Dr. Woodcock considered all relevant evidence in making her decision, the SDR Board does not recommend returning this matter to the Center Director for corrective action. We also believe that, for reasons discussed above, remanding this matter to the Center Director would be futile.

¹⁴¹ Woodcock Decisional Memorandum at 5-9.

¹⁴² *Id.* at 9.

¹⁴³ *Id.* at 10.

¹⁴⁴ Appeal at 20 (emphasis in original).

CONSIDERATIONS FROM THE ACTING CHIEF SCIENTIST

In my capacity as Acting Chief Scientist, I feel the responsibility to convey some comments regarding the underlying science for the decision being challenged by Dr. Unger in his appeal. I cannot begin to understand the depth of pain and suffering that patients with DMD and their families endure. As an experienced physician, I struggle to identify any other diseases associated with this degree of suffering, not only to patients but to their families. Nevertheless, my assessment is that the data presented by the sponsor to date are not adequate to support accelerated approval of eteplirsen

Studies in animals showing that eteplirsen leads to “exon 51 skipping” are an important first step in assessing whether eteplirsen might work for a subset of patients with DMD because skipping exon 51 is necessary for the production of dystrophin in these patients.¹⁴⁵ The next step is to assess whether eteplirsen actually leads to the production of dystrophin in patients with DMD and, if so, whether such an increase in dystrophin confers clinical benefit. Despite the promising animal studies demonstrating exon 51 skipping, both Drs. Woodcock and Unger, as well as the review team in CDER, agree that the amount of dystrophin produced in the clinical studies conducted at doses of up to 50mg/kg per week is very low. Animal data suggest that the doses studied in humans is too low; in animals, exon 51 skipping was detected in a nonlinear, dose-dependent manner (that is, higher doses led to significantly more exon 51 skipping). Specifically, 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. These dose-dependent responses are important because it is wholly conceivable that higher doses would lead to a much greater amount of dystrophin production, which could be important for clinical benefit. Because the drug appears to be safe, the review team recommended evaluation of much higher doses of eteplirsen, of at least 200mg/kg per week. Approving a drug at a dose that does not show a meaningful increase in dystrophin (when the drug could theoretically achieve one at higher doses) is concerning.

As for accelerated approval, the regulatory standard at issue requires a sponsor to show that the drug under review leads to an effect on the surrogate endpoint (in this case, the production of dystrophin) and that the effect is reasonably likely to predict clinical benefit (in this case, improving, or slowing down decline in, muscle function). The term “reasonably likely to predict” acknowledges the potential for doubt in the outcome of interest. Indeed, nobody knows the minimum level of dystrophin that is likely to confer clinical benefit in patients with DMD. The critical scientific and regulatory issue at stake in CDER’s decision here is whether such minute amounts of dystrophin are reasonably likely to predict clinical benefit at the dosage of the drug subject to approval. In this case, both Drs. Woodcock and Unger have attempted to provide a rationale, based on scientific and professional judgment, for whether or not such small levels of dystrophin are reasonably likely to predict the clinical effect of interest. By any meaningful objective standard, however, the overall evidence derived from eteplirsen’s limited clinical development program does not support that the levels of dystrophin produced by eteplirsen at the doses studied are reasonably likely to provide clinical benefit. As pointed out in Dr. Unger’s appeal, “Study 201 did not show a treatment effect on its 1^o clinical endpoint, change in 6-minute walk distance at Week 24. Study 202 failed on the same endpoint at 48 weeks. The course of these Study 201/202 patients, having received eteplirsen for some 3.5 years, was not distinguishable from external control patients.”¹⁴⁶

¹⁴⁵ Eteplirsen targets a subset of patients with DMD who are amenable to exon 51 skipping.

¹⁴⁶ Appeal at 16.

Some may argue that it would be reasonable to proceed with accelerated approval based on eteplirsén's safety profile, even where there are significant doubts about the drug's effectiveness. That argument does not take into account the risks of treatment with indwelling catheters to maintain vascular access in young patients, who would otherwise not need one and who often receive adjunct chronic corticosteroids, or, even more importantly, the detrimental impact on their quality of life.

I would be remiss if I did not note that the sponsor has exhibited serious irresponsibility by playing a role in publishing and promoting selective data during the development of this product. Not only was there a misleading published article with respect to the results of Study 201/202¹⁴⁷—which has never been retracted—but Sarepta also issued a press release relying on the misleading article and its findings. As determined by the review team, and as acknowledged by Dr. Woodcock, the article's scientific findings—with respect to the demonstrated effect of eteplirsén on both surrogate and clinical endpoints—do not withstand proper and objective analyses of the data. Sarepta's misleading communications led to unrealistic expectations and hope for DMD patients and their families. It is very disappointing that the findings did not hold up to careful review.

FDA must remain steadfast in its commitment to alleviating pain and suffering, approach the most challenging problems with absolute determination, and apply maximum flexibility to facilitate the development and availability of effective treatments. The agency's value centers on its ability to do all of the above while maintaining objectivity, even in the face of political pressure. FDA should never mislead patients by granting even accelerated approval to products that are not shown to offer the prospect of meaningful benefit to patients under the appropriate regulatory and scientific standard.

I acknowledge that there are currently no specific drugs available to treat patients with DMD and that issuance of a complete response letter would cast uncertainty on whether eteplirsén would continue to be developed, based on business and financial decisions that are external to FDA. However, approving products based on hope, on subjective clinical judgment, or on theoretical constructs that are not anchored in data leads to irreparable damage to patients. Approval at this time could deter others from pursuing the development of truly effective treatments, both for DMD and other serious, life-threatening conditions. Granting accelerated approval here on the basis of the data submitted could make matters worse for patients with no existing meaningful therapies—both by discouraging others from developing effective therapies for DMD and by encouraging other developers to seek approval for serious conditions before they have invested the time and research necessary to establish whether a product is likely to confer clinical benefit.

I remain deep in my conviction that, through science and a flexible, sound regulatory approach, good therapies will emerge to provide meaningful clinical benefit to patients with DMD and other rare serious diseases.

THE SDR BOARD'S ADDITIONAL RECOMMENDATION

Although the SDR Board acknowledges that the scope of our review, as prescribed by the SDR-SMG, is limited to procedural questions, we nonetheless feel duty-bound to make one additional recommendation. As noted above, Dr. Unger seeks from the Office of the Commissioner a substantive, scientific review of Dr. Woodcock's decision to grant accelerated approval to

¹⁴⁷ See Mendell JR, et al. *Ann Neurol* 2013;74:637-47.

eteplirsen. The SDR-SMG presumes that an initiator such as Dr. Unger has received some substantive review of the scientific dispute at issue as part of a formal appeals process in the Center. Dr. Unger has never received any such formal review of his scientific arguments or the underlying evidence. To the extent he has ever received any substantive review of his scientific disagreement with Dr. Woodcock, Dr. Woodcock herself was the one who conducted that review and resolved the conflict in her own favor. Neither the SDR-SMG nor CDER's SDR-SOPs contemplate a scientific disagreement that arises between a Center Director and another manager in that same Center—partly because no one has ever anticipated the unique circumstance of this case. Especially given the SDR Board's concerns regarding the decision-making process at CDER, we think additional review within the Office of the Commissioner is appropriate.

The SDR Board encourages you to conduct a thorough substantive review of the scientific dispute in this matter or, in the alternative, to convene a panel of relevant experts to conduct such a review and provide advice to the agency and you, as Commissioner, on whether the evidence of the effect of eteplirsen on the surrogate endpoint is reasonably likely to predict clinical benefit. If you choose the latter, in light of the public and political pressure evident during the entire review process at CDER, as detailed in this recommendation, we believe that delegating this critical evaluation to a panel of experts would help ensure that the agency makes the most appropriate decision from the perspective of protecting patients and the public health, especially for DMD patients. Knowing as we do that you value cross-Center collaboration with respect to medical product development, we recommend that you include on the panel experts from other Centers devoted to the regulation of medical products. Doing so would not only help ensure diverse expertise on the panel but also provide insights on the effects that any proposed regulatory decision on eteplirsen might have on products regulated by those other Centers. We further recommend that you consider whether to include experts from other components within the Department of Health and Human Services and whether, consistent with applicable laws and the appropriate timeframe for a decision, you should also include outside experts on the panel.

Agency Scientific Dispute – Appeal

Date: July 18, 2016

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

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

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

1. Background

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

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

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

Disease Background:

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

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

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

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

Drug Background:

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

Drug Development Background:

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

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

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

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

a. Clinical/Statistical Efficacy

Accelerated Approval:

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

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

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

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

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

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

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

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

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

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

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

1. Production of an appropriate mRNA transcript

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

2. Quantitative assessment of dystrophin content in muscle biopsies by Western blot

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

a. Study 201:

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

b. Study 202:

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

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

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

1. The goal was to assess the change in dystrophin with treatment, i.e., pre-treatment vs. post-treatment, but most of the baseline biopsies were obtained from subjects external to Study 201/202, who could differ in unknown ways from subjects in Study 201/202.
2. For all patients, the Week 180 biopsies were obtained from different muscles than the baseline biopsies, and studies of both normal human muscle and non-clinical DMD models have shown that dystrophin levels vary among muscles.
3. The baseline biopsies for the three subjects with Week 180 data had been stored for several years and the protein may have degraded, leading to a falsely low baseline value, and a greater apparent increase from baseline, accordingly.

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

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

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

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

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

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

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

c) Study 301:

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

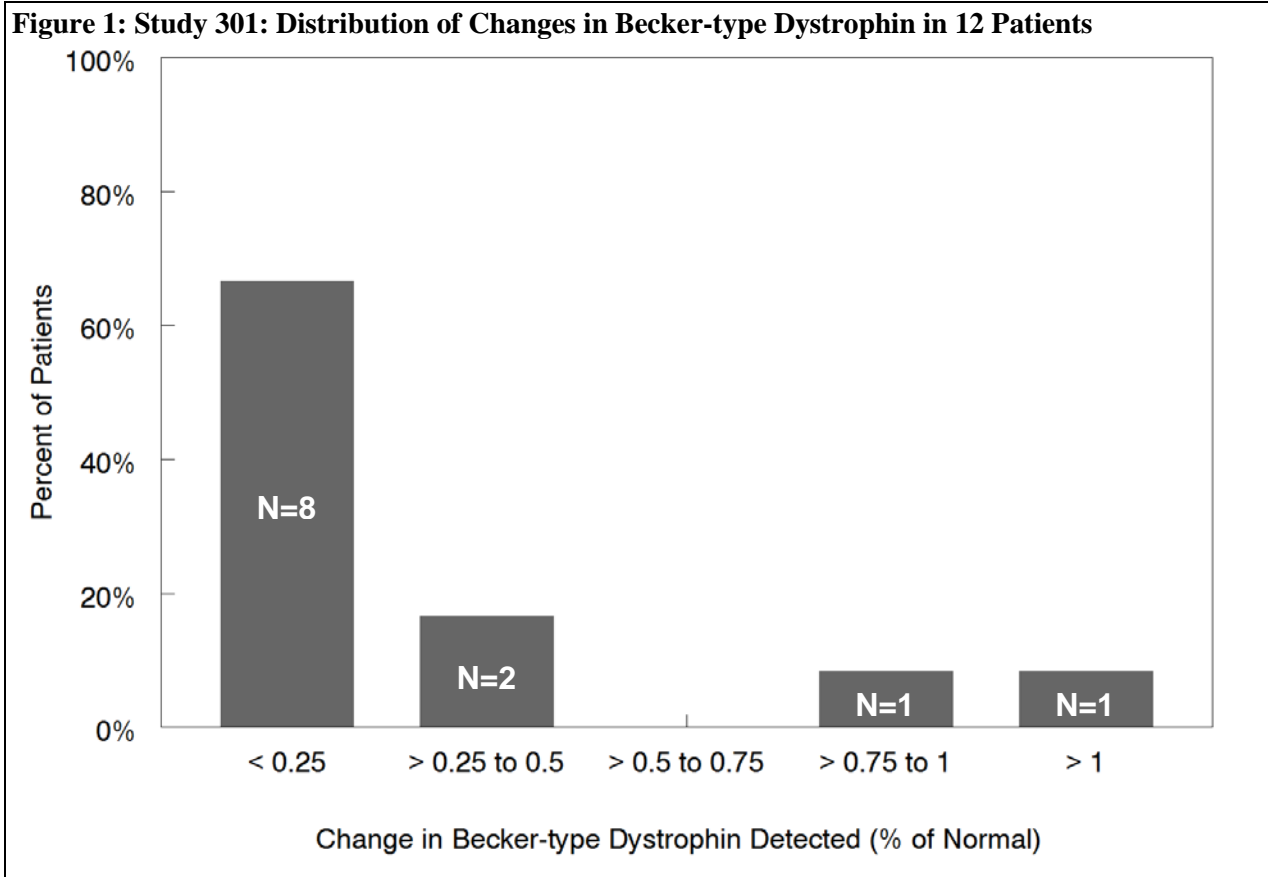
The data are shown in Table 1 and the distribution of these changes is shown graphically in Figure 1. Of these 12 patients, 8 (two-thirds) had a change of 0.25% or less; only 1 patient (8%) had a change greater than 1%. The applicant used 3 methods to consider the numerous values below the limit of quantification,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

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

Patient	Time	status	value (%)	mean (%)	delta (%)	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		
		pass	0.29					fail	0.05		
2	Baseline	pass	0.35	0.35	0.01	9	Baseline	fail	0.14	1.57	1.33
		fail	0.26					pass	0.24		
	Week 48	pass	0.36	0.36			Week 48	fail	1.17		
		fail	0.12					pass	1.57		
3	Baseline	pass	0.06	0.06	0.31	10	Baseline	pass	0.11	0.11	0.01
		pass	0.06					fail	0.05		
	Week 48	pass	0.5	0.37			Week 48	pass	0.12		
		pass	0.24					fail	0.11		
4	Baseline	pass	0.04	0.04	0.06	11	Baseline	pass	0.01	0.05	0.43
		fail	0.06					pass	0.08		
	Week 48	pass	0.1	0.1			Week 48	pass	0.31		
		fail	0.19					pass	0.63		
5	Baseline	fail	0.1	0.17	0.85	12	Baseline	pass	0.02	0.02	0.07
		pass	0.17					fail	0		
	Week 48	fail	0.92	1.02			Week 48	pass	0.09		
		pass	1.02					fail	0.01		
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		
		fail	0.29					pass	0.21		
7	Baseline	fail	0.04	0.17	0.25	13	Baseline	fail	0.34	0.21	0.03
		pass	0.17					pass	0.21		
	Week 48	fail	0.22	0.42			Week 48	fail	0.34		
		pass	0.42					pass	0.21		

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

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



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

Study 201/202 – Data through Week 48

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

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

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

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

Study 201/202 – Week 180 Data

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

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

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

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

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

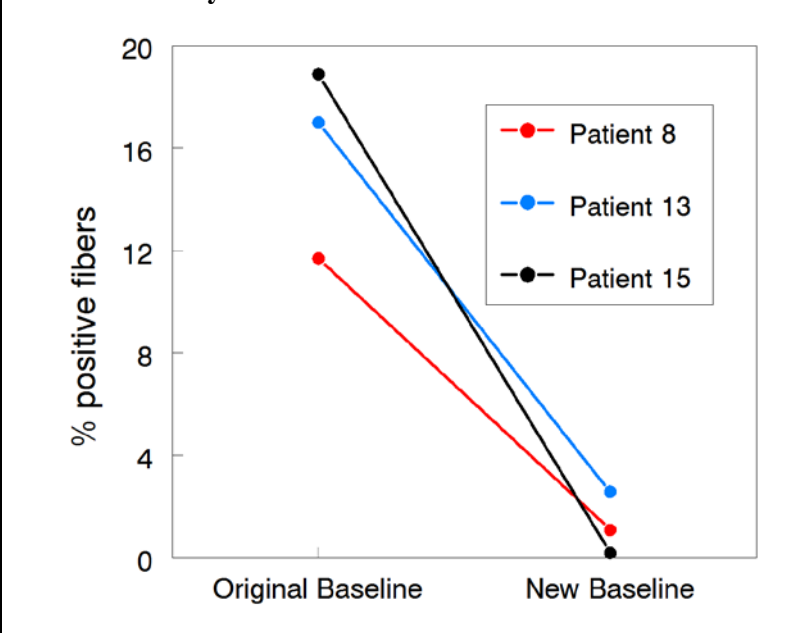
As noted above, there were 3 patients in Study 201/202 with adequate archived tissue from baseline, which permitted a new immunohistochemistry analysis and a comparison of results between the old and new methods. Figure 2 shows how the two methods compare.

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

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

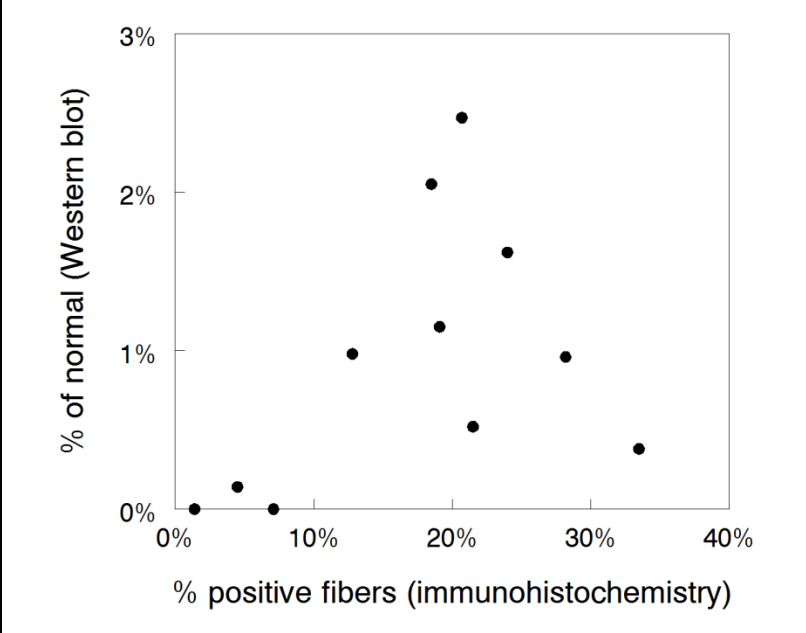
The integrity of the applicant's data is further called into question by lack of agreement between the immunohistochemistry and Western blot methods, i.e., a lack of internal consistency. The applicant claims to have enhanced both the immunohistochemistry methods and the Western blot methods in preparation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

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



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

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



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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

- Dystrophin levels >10% on Western blot are usually associated with a BMD phenotype. Within the BMD phenotype, the relation between disease severity and protein expression is not clear. Protein quality, rather than quantity, may play a key role in determining phenotype in BMD.
- Patients with DMD are usually found to have undetectable levels of dystrophin, or very low levels. Dr. Woodcock notes that she believes the conventional threshold of <10% protein resulting in DMD was based on immunohistochemistry data. She tries to make a conversion between values observed from immunohistochemistry (~10% points higher on immunohistochemistry than Western blot in DMD) and those observed from Western blot, but I caution that immunohistochemistry results, in particular, are highly method-dependent, as noted above.
- Rarely, dystrophin levels in the 3 to 10% range have been associated with Becker Muscular Dystrophy phenotypes. Dr. Woodcock found no evidence of a threshold value for protein content and expression of a DMD phenotype.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

2) *The effect size is inadequate on its face.*

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

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

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

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

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

“The broad phenotypic distinctions made in the clinic (e.g., DMD vs IMD vs BMD) are different from the prediction of benefit to an individual patient who has a specific baseline dystrophin level and whose mutation and external factors do not change pre- and post-drug. For example, extending ambulation by six months to a year would not normally move a patient from one to another of these categories, but could be very important to quality of life (e.g., as suggested in the Bello study). This is also true for other functional improvements.

For these reasons, incorporating the analysis of dystrophin content discussed above, I conclude that the biochemical data strongly support the idea that low-level increases in dystrophin production are reasonably likely to predict clinical benefit.”

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

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

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

Dr. Woodcock states:

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

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

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

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

Correlation between dystrophin production and clinical effect

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

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

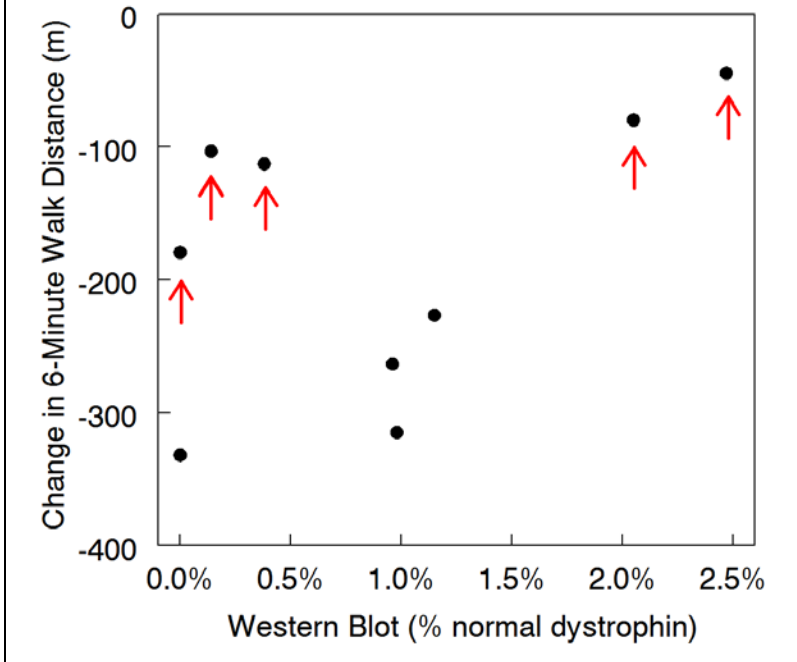
The Division's view notwithstanding, it is worth considering patients on an individual basis to assess the correlation between the quantity of Becker-type dystrophin detected and changes in physical performance.

As noted by Dr. Woodcock, the 6-minute walk test results do not show a strong correlation (Figure 4). For the 9 patients in Study 201/202 who remained ambulatory at Week 180 and agreed to undergo a fourth muscle biopsy, the figure shows little correlation between the quantity of dystrophin detected (x-axis) and preservation of physical function as assessed by the change in 6-minute walk distance from baseline (y-axis) after weekly infusions of eteplirsen for 3 to 3.5 years.

For the 5 patients whose 6-minute walk performance was best preserved (red arrows), 2 had the highest dystrophin levels detected in the study (upper right), but 3 had levels that were near-zero (upper left).

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

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



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

Given that the correlation was driven by the patient depicted at the lower right (blue arrow; dystrophin level =

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

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

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

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

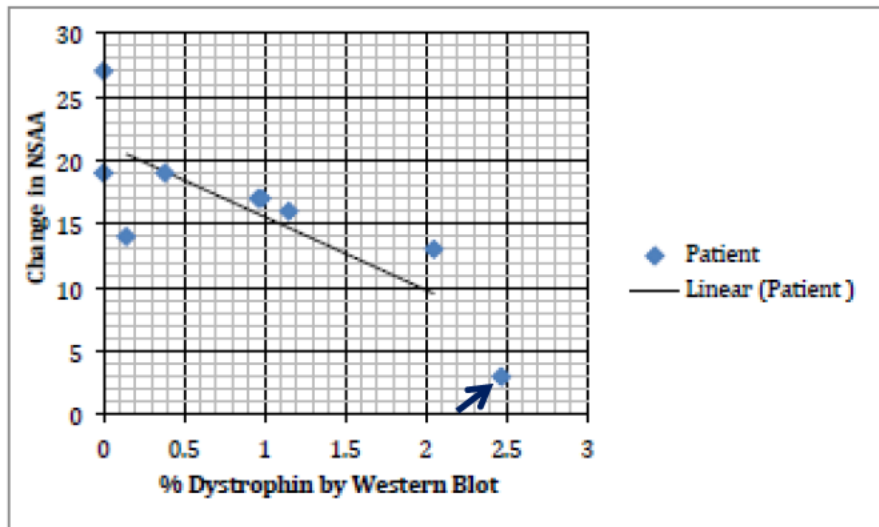
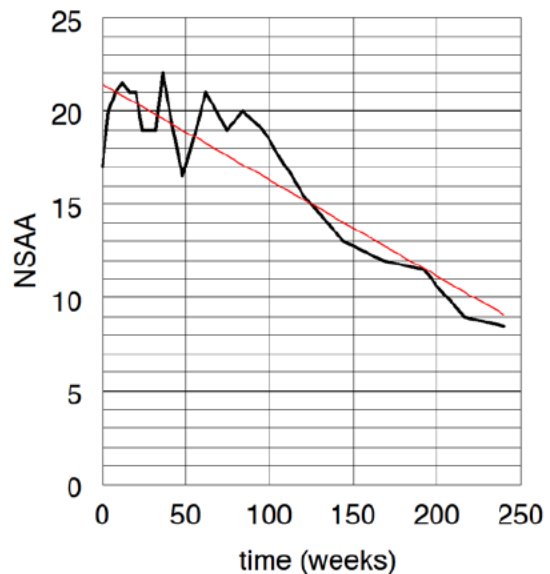


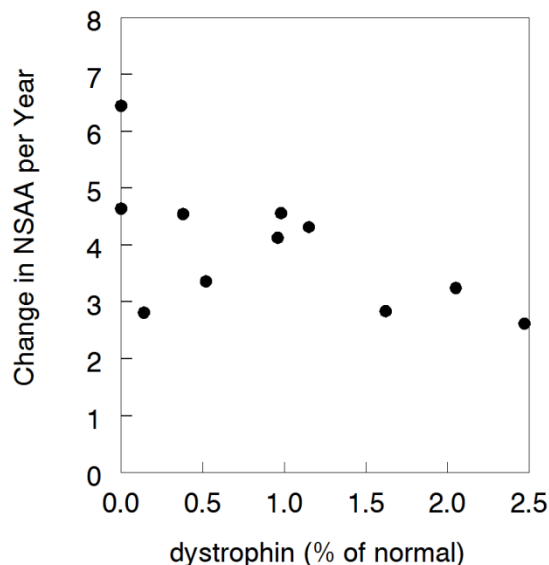
Figure 6: Patient 006: NSAA vs Time



Summary:

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

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



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

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

Dr. Woodcock concludes:

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

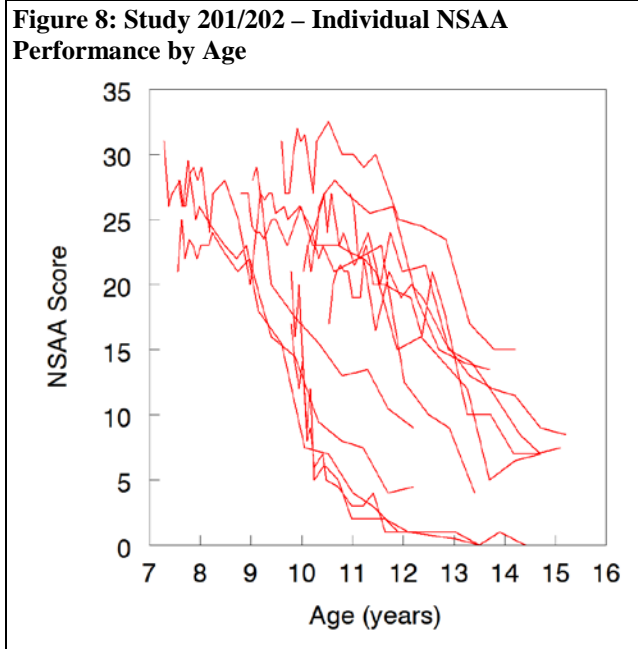
alternative therapies, while preserving standards for safety and effectiveness. For drugs granted accelerated approval, postmarketing confirmatory trials are required to verify and describe the anticipated clinical benefit, and FDA may withdraw approval of a drug if a trial required for verification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

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

4. With the false perception that eteplirsen is effective, patients who are gaining benefit from steroids but experiencing untoward side effects might be inclined to taper or stop them, which could lead to more rapid disease progression.
5. False scientific conclusions have the potential to mislead the field of medicine, slowing progress in finding and developing therapies that actually *are* effective. For example, consider the scenario of a related drug with far greater potential to promote dystrophin production in patients with DMD. In order for a sponsor to study such a drug, patients would likely have to agree to discontinue eteplirsen, and few patients may be willing to do so. In short, approval of an ineffective therapy has the potential to discourage or inhibit the development of other drugs that *are* effective, and this impact can be significant.
6. Accelerated approval would lower the evidentiary standard for effectiveness to an unprecedented nadir. The amount of dystrophin produced in Study 301 is so meager that it could be considered to be tantamount to any increase in dystrophin. In other words, if a statistically significant change of 0.3% – a mere 3 parts out of a thousand – is considered adequate to support accelerated approval here, then the question arises as to whether there would be any statistically significant change that would be too small to be considered “reasonably likely” to support accelerated approval. Similarly, if a ‘responder’ had been defined as a patient with an increase in dystrophin of $\geq 1\%$ (and there is no basis to accept such a low threshold), there would have been only a single responder in Study 301. If we were to adopt the concept that, for rare diseases, accelerated approval could be supported by any statistically significant change in an appropriate surrogate, or a response in a single patient, we would enable accelerated approval of a myriad of drugs for rare diseases. No doubt there are some who would applaud this as an advance. But a standard this low would undercut FDA’s ability to ensure that drugs that are approved are effective; it would call into question much of what we do. Lowering the bar to this level would be tantamount to rolling back the 1962 Kefauver-Harris Drug Amendments to the Federal Food, Drug and Cosmetic (FD&C) Act, which have served Americans well for some 54 years.
7. With accelerated approval of this NDA, there would be highly detrimental effects on drug development. Traditional drug development for rare diseases might be replaced by a system where small, baseline-controlled, proof-of-concept studies designed to show any change in a surrogate marker would provide a basis for accelerated approval, assuming that the pathogenesis of the disease was well understood and that the surrogate was directly on the causal path. There would be little reason to pursue adequately controlled clinical trials to support efficacy prior to accelerated approval; in fact, the possibility of

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

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



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

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

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

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

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

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

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

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

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

5/31/2016	Center Director Briefing	Discuss reviews conducted by the review team and leadership along with any additional information obtained from the sponsor. Discussed Dr. Woodcock's memo. Timeline for reviews due to Dr. Woodcock.
7/6/2016	Center Director Briefing	<ol style="list-style-type: none"> 1. The levels of dystrophin observed in 12 DMD patients from the recent interim analysis of an ongoing trial and whether the levels seen can be interpreted to be "reasonably likely to predict clinical benefit" and used as a surrogate endpoint to support accelerated approval. 2. The design of one or more PMR trials to confirm clinical benefit of eteplirsen if it is approved under accelerated approval. 3. Description of the available clinical data in the drug label if approved.

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

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

5. Action, Decision or Remedy Sought

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

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

6. Path Forward

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

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

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

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

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

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

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

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

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

CENTER DIRECTOR DECISIONAL MEMO

NDA# **206488**

Drug Name **EXONDYS 51 (eteplirsen)**

Indication **Duchenne Muscular Dystrophy (DMD)**

Sponsor **Sarepta**

Author **Janet Woodcock, M.D.**

**Director, Center for Drug Evaluation and Research (CDER),
Food and Drug Administration**

SUMMARY

This memorandum explains the CDER’s final decision on the above application. I have read the reviews and recommendations by Drs. Unger (Office level), Bastings (Division level), Farkas (Cross-Discipline Team Lead), Breder and Rao (Clinical Reviewers), Ling (Statistical Reviewer), and Bhattaram, Wu, and Rogers (Clinical Pharmacology Reviewers). In addition to the review memoranda, I have also reviewed the Advisory Committee briefing materials, pertinent portions of the sponsor’s submission, and multiple scientific statements submitted by the public, including a letter from a large number of DMD experts.

The review team has done an exemplary job in performing a detailed evaluation of the data submitted with the application. Nevertheless, I disagree with certain of their findings and come to a different conclusion, as discussed below.

I find that the data contained in NDA 206488 meet the standard for accelerated approval under 21 CFR 314. 510 based on the surrogate endpoint of increased dystrophin protein production, a surrogate endpoint that I conclude is reasonably likely to predict clinical benefit.

DISCUSSION

Extensive analyses have been performed by the team on the clinical results of the long-term experience of 12 patients administered the drug, and I will not recapitulate these.

Approval under 314.510 is based, among other things, on adequate and well-controlled clinical trials establishing that the drug product has an effect on a surrogate endpoint that is reasonably likely to predict clinical benefit. Below, I discuss how both of parts of this standard are met.

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

The characteristics of adequate and well-controlled studies are laid out in 21 CFR 314.126. Three lines of evidence are pertinent to the conclusion that eteplirsen results in increased dystrophin production.

- Production of an appropriate mRNA transcript
- Quantitative assessment of dystrophin content in muscle biopsies by Western blot
- Semi-quantitative assessment of dystrophin in muscle tissue by immunohistochemistry (IHC) techniques

The sponsor provided data demonstrating an increase in mRNA expression following treatment with eteplirsen. The drug's proposed mechanism of action is to bridge a section of the pre-RNA to result in a shorter mRNA with an open reading frame, e.g., "exon skipping." In this case, the production of an appropriate mRNA transcript has been documented by PCR and Sanger sequencing. Although this establishes proof of mechanism, it does not mean that there is increased protein production.

In the following, I discuss the assessments related to dystrophin protein production (2. and 3.) in some detail. Much of the controversy over the adequacy of these assessments relates to the fact that rigorously validated assays were not used to evaluate the initial 3 muscle biopsies, apparently resulting in overestimation of the various readouts and some irreproducibility of IHC and Western blot dystrophin assays. For these reasons, I do not discuss or rely upon the results of these earlier assays, or on re-reads of them. With FDA's assistance, the sponsor improved the design and conduct of the assays and performed repeat biopsies on 11 of 12 patients at week 180. The control samples for these week 180 biopsies were stored baseline tissue (in 3 of 11 subjects) and baseline biopsies from subjects with exon 51 amenable mutations enrolled in another trial by the sponsor. FDA reviewers had the following concerns about these controls, leading them to conclude that the studies were not adequate and well controlled.

1. Most of the baseline biopsies were not from the same subjects as the week 180 biopsies (as the original tissue had been used up for the previous assays). Given this, the control subjects could differ in unknown ways from the test subjects.
2. The biopsies taken at week 180 were from different muscles in the upper extremity than the baseline biopsies, including subjects with baseline tissue as well as for control samples. It is hypothesized that there may be differences in dystrophin protein content among various muscles in DMD patients.
3. The existing baseline biopsies for the three subjects with 180 week data had been stored frozen for several years and may have changed (apparent decrease in dystrophin protein content) over time.

In my judgment, these issues increase the uncertainty around the results, but do not necessarily render them an inadequate basis on which to draw a conclusion. The non-treated control subjects were very similar in age and dystrophin mutation site to the treated subjects (sponsor Appendix 10, AC briefing package). The single deltoid muscle biopsy in the untreated control group (subject 7, sponsor Appendix 14, AC briefing package) had replicate dystrophin levels of 0.3% and below the limit of quantification, averaging out at below 0.3%, and not different than biceps biopsy results in other patients, suggesting

that variations in upper extremity biopsy site (concern b above) did not result in large differences in the findings. There was little difference in the dystrophin protein content found in the stored baseline samples and the frozen samples, as discussed below.

The data submitted with the original application, supporting the finding that eteplirsen increases the production of dystrophin protein, come from the quantitative assessment of (internally truncated) dystrophin in muscle tissue by Western blot using the controls described above. Much of the controversy around this method relates to the fact that the apparently achieved dystrophin levels are very much lower than originally hoped (and previously claimed by the sponsor and investigators).

In the 180 week assessment, the three subjects with baseline biopsies available had baseline dystrophin levels (reported as % of normal) below the level of quantification of the assay used (0.25%). These results were similar in magnitude to the baselines of the six additional control biopsies drawn from subjects in another study (highest level 0.37%). At week 180, two treated subjects had (an average of replicate) dystrophin levels above 2%, two had over 1%, and two additional had about 1%. Of these individuals, two subjects having both baseline and week 180 samples had clearly increased levels at week 180 compared to baseline. (The third subject with a baseline sample did not consent to a week 180 biopsy). Unsurprisingly, some subjects had week 180 dystrophin levels similar to the overall baseline control levels. Not all individuals are expected to respond to a drug intervention. The issue is whether the dystrophin levels found at 180 weeks were within the variability expected for this assay in such patients and, thus, could have arisen by chance, or whether they could have been caused by differences from the controls or from sample storage as outlined above, or whether they reflected a drug effect, and, thus, whether these data could be seen as adequate and well-controlled. The following data are relevant to this issue.

Because the original data on the presence of dystrophin by Western blot suffered some difficulties in interpretation because of lack of availability of baseline samples from most patients, the sponsor of this application submitted, subsequent to the Advisory Committee meeting on this drug, additional Western blot data from 12 patients with baseline and 48 week eteplirsen exposure, using baseline and post-treatment muscle biopsies from the same patients and muscle groups. This experiment clearly shows, using adequate controls, that the drug increases dystrophin protein production in some of the patients. The mean baseline dystrophin values in this study were very similar to the mean baseline values in the 180 week study. The achieved levels of dystrophin in these patients are lower than those seen in the Western blots from the week 180 patients. Only 2 of 12 patients achieved a level over 1% of normal control. It is not known if this result is due to a shorter duration of drug exposure or to other factors. Putting together the 180 week data and the additional 48 week data, I conclude that there is substantial evidence from Western blot experiments of increased dystrophin protein production, albeit at a low level.

A finding of increased dystrophin was also seen in several IHC assays performed by the sponsor. Both assays were originally performed with baseline and several pre-180 week assays by the sponsor as a part of the clinical trial. The validity of the results of these assays were questioned by FDA because of methodological problems in their conduct, as documented in the primary clinical review and in the inspection report. Therefore, I will not further consider the results of these original assays. As discussed for the Western blot above, the sponsor responded by performing an additional 180 week biopsy and repeating the assays. Baseline tissue was available, as for Western blot, from recut samples

in only three cases. In one of these, the subject did not consent to a biopsy at 180 weeks. To supplement the three baseline samples the sponsor included six other untreated patients from a different trial, as discussed above for the Western blot. In both assays, greater staining or intensity was observed after drug exposure at week 180 compared to controls. The results are described in more detail below.

A Percent Dystrophin Positive Fibers analysis was a semi-automated evaluation performed at 180 weeks and compared to the controls used for the 180 week study as discussed above. The percentage of positive fibers was assessed using a blinded read by Nationwide Children's Hospital and by three independent pathologists through Flagship Biosciences. The technique used to assess percent positive fibers was modified from the original assay in the following ways:

1. A computer algorithm (MuscleMap from Flagship) that performs non-linear mapping of all fibers was used for consistent and automated analysis of low intensity values, in contrast to a manual and non-standardized fiber counting technique in the prior assay.
2. The images were inverted and amplified to score the total fibers (the denominator for the percent positive fiber scoring).
3. An isotype matched secondary antibody staining step was incorporated to confirm lack of non-specific staining and reduce background noise. The background signal was subtracted from test sample values in calculation of percent intensity.
4. 8% of the images for re-analysis were blinded, renamed, randomized, and rotated 180 degrees.
5. A rejection factor for the inter-rater analysis score of <4 was established.
6. The images were acquired in a more systematic and random fashion to minimize bias, with predefined rules for random sampling of fields and avoiding artifacts.

These changes were likely to result in a more conservative reading of Percent Dystrophin Positive Fibers, and indeed the results, including the new untreated baseline controls, were read at 1.1% positive fibers (in contrast to a higher result in the prior baseline using the original technique). The 180 week cohort had a score, using this technique, of 17.4% positive fibers, showing a statistically significant difference. Now, these results are subject to the same caveats as discussed for the Western blot (1-3 above), in that there were only two baseline to 180 week pairs, that the baseline samples had been frozen for years, and that the external controls might differ in some way. So, these results cannot stand alone.

Other reviewers have pointed out that the (much higher) baseline values for Percent Positive Fibers from the original experiment are not very different from the 180 week values in this new experiment. However, I would point out that experimental conditions changed quite a bit, and very low values for all the external controls, statistically comparable to the frozen baseline results, were obtained in this recent experiment, suggesting that it returned a more conservative result. I do not believe that comparison of the original baseline data, obtained under one set of experimental conditions, can be compared to the later 180 week results, done under different, more optimized conditions and yielding very different results for new (external control) baseline samples.

The sponsor also performed a Mean Relative Fluorescence Intensity assay for dystrophin. This assay is commonly performed by laboratories evaluating DMD patients and is intended to be a semi-quantitative evaluation of dystrophin content. Using the six external baseline samples and the three stored study patient baseline samples, the mean intensity approximately doubled from baseline to 180 weeks. The technique for this assay did not change significantly from the technique used in the assay done as part of

the original protocol, and the baseline means for the patient samples were roughly comparable to the baseline means obtained in the new experiment.

Although the IHC assays provide only semi-quantitative assessments of dystrophin content, they do support an effect of eteplirsen on the proposed surrogate endpoint (an increase of dystrophin production as a result of drug exposure). The accompanying microscopy images also demonstrate correct localization of the molecule within the muscle fibers, a very important factor in any translation to clinical benefit.

In summary, I conclude that there is evidence from adequate and well-controlled trials, and supportive evidence, that exposure to eteplirsen increases dystrophin protein production in muscle cells.

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

In this case, the standard for clinical benefit does not require “cure” or “conversion to Becker MD (BMD) phenotype.” Clinical benefit encompasses improvements (including slowing of disease progression) in how an individual feels or functions, or an improvement in survival. There is no question that, for DMD patients and their families, small improvements in function or delays in loss of function are meaningful benefits. Therefore, the question is:

What amount of increase in dystrophin production is reasonably likely to predict clinical benefit (even small benefits)?

The usual way to address this question would be to rigorously evaluate what is known about the correlation between dystrophin levels in muscle and expression of disease. The following summarizes the existing scientific literature on this topic and the challenges in interpreting it.

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

Experts usually classify patients clinically as DMD (severely affected at a young age); intermediate MD (also called DMD/BMD); or BMD, which can range from severe BMD to asymptomatic individuals with biochemical abnormalities, usually increased creatine phosphokinase (CPK). There is clearly a wide spectrum of disease wherein the ends of the spectrum are easily distinguishable, but the zone of real interest for this discussion, between DMD and intermediate presentations, is not rigorously categorized. In part, this is because “intermediate muscular dystrophy” (IMD) is less common, due to the consequences of having either in-frame mutations with a truncated protein expressed (leading to BMD) or out-of-frame mutations with little-to-zero protein expressed (leading to DMD), as discussed below.

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

Anthony, et al., (*Neurology*, 83, 2014) in a collaborative cross-laboratory study, investigated the variability of techniques used to quantify dystrophin in individuals with muscular dystrophy. Blinded tissue sections from three DMD and three BMD muscle biopsies were tested in five

different laboratories accustomed to performing dystrophin quantification. Estimates of dystrophin expression using a somewhat standardized IHC technique were about 20%, 11% and 10% of normal for the three DMD samples, on average among the laboratories. Corresponding estimates of dystrophin content by Western blot, using an actin antibody to normalize for loading, but not a serially diluted standard control, resulted in dystrophin estimates of about 11%, 0, and 0.4% respectively, with fairly high CV's. Therefore, in this small sample, repeated across five experienced laboratories, IHC estimates were about 10 percentage points higher than Western blot estimates.

Significantly higher estimates by IHC by fluorescence intensity (overall about 23% of normal) than by Western blot were also seen in the evaluation of week 180 muscle biopsies in the Sarepta trial. Because much of the historical data on protein content vs phenotype has been reported using IHC analysis, extrapolating these findings to the current trial data is challenging. Additionally, Anthony et al., found that the inter-laboratory variability was greatest for the low levels of dystrophin found in the DMD patients. Western blot data in the literature quantifying dystrophin and relating it to phenotype is often from experiments that were not designed to distinguish among dystrophin levels below 10% of normal. These may have been reported out as "less than 10%." From this sponsor's well-controlled studies, the analytically accurate dystrophin baseline for many DMD patients might be in the range of 0.02-0.35 % normal, hence previous estimates of 5-10% might be an over-estimation using non-standardized and semi-quantitative methods.

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

Significantly, if the epitope recognized by the antibody is modified by the deletion, the dystrophin isoform may not be recognized and a result read out as zero. For this reason, recent studies use multiple antibodies against known regions. Additionally, muscle biopsies in patients with BMD and DMD may be quite variable in degree of fibrosis and fatty replacement; this may decrease the reproducibility and representativeness of muscle biopsy estimates of dystrophin content by Western blot. Additionally, imaging methods, choices for normalization, biopsy handling, background standing, and a multitude of other experimental conditions can influence results.

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

Dystrophin is a very large protein with multiple functional domains. Generally, DMD results from an out-of-frame mutation (often a deletion) that leads to an unstable or unreadable mRNA transcript. Thus, DMD patients usually have zero or very low levels of dystrophin, but the DMD phenotype can also result from in-frame mutations that result in a unstable transcript or dysfunctional dystrophin isoform. BMD usually results from an in-frame mutation (often an exon deletion) that affects the functional quality of the protein and also the quantity produced. It remains unclear what role protein function plays vs quantity in leading to the wide range of variability in BMD phenotypes. There are a vast number of mutations that can lead to each of these phenotypes (Tuffery-Giraud, et al., *Hum Mutat*, 30, 2009), all of which can have different effects on protein function as well as protein production. This micro-heterogeneity is common in genetic diseases and is highly germane to

evaluation of interventions targeting the gene, gene expression, or protein function. There are also non-dystrophin-related factors that can modulate phenotype.

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

I note that in the decades since 1988, much technical progress has been made in standardizing Western blot techniques, and the results from early studies may not be fully comparable to those from recent experiments.

- a. The seminal 1988 paper on this subject (Hoffman et al., *NEJM*, 318(21)) found that the majority of patients with DMD had undetectable levels of dystrophin using their Western blot technique and that 35 of 38 had levels below 3% in their assay. They also reported that one of seven “intermediate” patients had dystrophin levels below 3% of normal, as did one of the 18 patients with a BMD phenotype.
- b. Beggs et al., (*Am J Hum Genet*, 49, 1991) published one of the early studies on the correlation between the level of dystrophin on Western blot and clinical features of BMD. Western blot was performed using a polyclonal serum and had about a 20% variability between blots according to the authors. In this study a number of patients with BMD or intermediate phenotype (DMD/BMD) were found to have dystrophin contents that overlapped with those of the DMD patients. Of four patients included with DMD phenotype, two had less than 5% dystrophin, and two had 10%, by their assay. Of patients with BMD/DMD phenotypes, eight were found to have 10% of normal dystrophin, two had 15%, one had 50%, and one had 100%. Three BMD patients with dystrophin levels of 10% were found; two of these had relatively mild disease.
- c. Nicholson et al., (*J Med Genet*, 30, 1993) studied patients across a wide range of DMD and BMD phenotypes. They used loss of ambulation as a criterion to establish five functional groups, grouped from one (most severe, LOA before age 9) to five (LOA past age 40) (pre-steroid era). *They found a linear relationship overall between dystrophin levels (Western blot with Dy4/6D3 antibody, using myosin for a loading control) and their five categories, with more dystrophin protein translating to better function. They found no significant difference between any two adjacent groups however, which they interpreted as showing considerable overlap, as reflected in their patient level data (Appendix 1), which showed a number of less severe patients (e.g., Group 2 or 3) registering no or very low dystrophin abundance on their Western blot assay.* Of note, they reported a higher average level of dystrophin protein in severe DMD patients than other investigators, partly resulting from 5 of their 21 severe patients reported to have dystrophin protein levels above 20.
- d. Neri et al., (*Neuromuscular Disorder* 17, 2007) reported on families with X-linked Dilated Cardiomyopathy. In these families, mutations give rise to absent dystrophin in heart muscle, but only reduced levels of nearly normal dystrophin in muscle tissue. One patient in their series had a normal neurological exam at age 23, an elevated CPK, and 29% of normal dystrophin protein in skeletal muscle by Western blot. This example can contribute to understanding the role of abundance of dystrophin protein vs compromised function.

- e. Anthony et al., (*JAMA Neurology*, 71, 2014) evaluated the correlation between phenotype and mRNA and protein expression in patients with both in-frame and out-of-frame mutations amenable to exon 44 or 45 skipping. Studying a group of patients with closely related deletions could diminish variability due to differences in function of the truncated protein. Five samples from patients with clinical “mild” BMD and in-frame mutations underwent Western blot analysis using the Dys-2 antibody. Their mean protein expression was 17% (normalized to actin) with a standard deviation of 7.5%. Two of the “mild” patients had dystrophin levels in this assay of around 10%. Based on comparisons of IHC experiments with various antibodies, the authors found “*no clear correlation between the level of dystrophin transcript or protein expression with clinical severity*” in 13 patients with in-frame mutations leading to BMD. The finding of Neri et al., above, along with this report, reinforce the concept that protein function (i.e., quality) is an important determinant of clinical severity and undermine the concept that 10% dystrophin protein content is a threshold, since these patients had “mild” BMD.
- f. Van den Bergen et al., (*J Neurol Neurosurg Psychiatry*, 85, 2014) compared dystrophin levels by Western blot with clinical severity in 27 patients with a clinical diagnosis of BMD. Dystrophin expression ranged from 4-71% and 3-78%, depending on the antibody used. *The authors found no linear relationship between dystrophin expression by Western blot using newly acquired muscle biopsies and clinical severity, muscle strength, or fatty infiltration on MRI.* Although this was the case for the majority of the patients, who had dystrophin levels above 20% of normal, four patients had levels at or below 10%. These patients generally had a more severe phenotype: one patient with a dystrophin level of 10% was wheelchair dependent at 45 years; one patient with a level of 7% developed trouble with stair walking at age 21; one patient with a level of 4% had a DMD phenotype with wheelchair dependency at age 10, one patient with a level of 3% had wheelchair dependency at age 25.
- g. Anthony et al., (*Brain*, 134, 2011) studied 17 BMD patients with exon 51 or 53 skipping-amenable mutations by IHC methods. These patients primarily had very mild or asymptomatic disease; the one patient classified as severe was ambulatory at age 25 but unable to run. *There was a statistically significant difference in dystrophin expression by IHC when patients classified as mild disease were compared to asymptomatic patients.*
- h. Bello et al., (*Neurology* 87, 2016) published a detailed study of loss of ambulation in DMD patients with particular exon deletions, using the CINRG-DNHS, a prospective natural history study. They found patients with exon 44 amenable mutations to have a two-year delay in loss of ambulation compared to the overall comparison group. This finding had previously been reported by another group (van den Bergen, et al., *J Neuromuscul Dis*, 1, 2014). The mutations studied (primarily single-exon deletion of exon 45) are known to undergo spontaneous skipping with production of some dystrophin. According to the Bello report, of six patients previously tested by IHC, three showed traces of dystrophin production and 0/four (possibly other patients) had dystrophin detectable by Western blot. These authors suggest that the observed differences in loss of ambulation (LOA) could be due to small amounts of spontaneously induced dystrophin that slightly ameliorate the ordinary DMD phenotype.

- i. Cirak et al., (*Lancet*, 378, 2011) published a study (AVI-4658) using intravenously administered eteplirsen that showed a detectable increase in dystrophin protein levels using both Western blot and immunofluorescence in 3/19 patients. The authors reported that the functional properties of restored dystrophin were confirmed by assessing increased levels and co-localization of neuronal nitric oxide synthase (nNOS) and sarcoglycan with dystrophin. Such a protein assembly is suggested to be indicative of functional restoration of the dystrophin-associated glycoprotein complex in muscle fibers (Molza et al., *JBC*, 290, 2015; Wells KE et al., *Neuromuscul Disord*, 2003). Cirak et al., reported that the restoration was more so in patients with exon 49-50 deletions than in those with 45-50 deletions, which is consistent with a previous observation that nNOS binding domain is located in dystrophin exons 42-45 (Lai Y et al., *J Clin Invest*, 2009). These studies suggest that important functional domains are included in the dystrophin protein induced by eteplirsen.

To summarize what is known about the association between dystrophin levels and phenotype, dystrophin content above about 10% on Western blot is usually associated with a BMD phenotype, except in patients with higher levels of dystrophin (including above 50%) who potentially have functionally deficient protein leading to a DMD phenotype. Within the BMD phenotype, a proportional inverse relationship between disease severity and protein expression has not generally been demonstrated (i.e., between 10-100%), although there may be a broad association, as seen in the Anthony study (*Brain*, 134, 2011). This may be due to the fact that protein quality, rather than quantity, plays a key role in determining phenotype in BMD. Patients with DMD are usually found to have no detectable, or very low levels of, dystrophin. Dystrophin content in the 3-10% range has been associated with DMD, DMD/BMD, and BMD phenotypes. I find no evidence of a threshold value for protein content and expression of a DMD phenotype, although the majority of DMD patients reported in the literature have dystrophin that is undetectable by the Western blot assays used. Generally, the divide between DMD and BMD, in terms of protein, is the result of the consequences of an OOF or an in-frame mutation, respectively. I believe that the conventional threshold, at or below 10% protein, was derived from the IHC data that seem to estimate low-level protein content about 10% percentage points higher on IHC than on Western blot, so that the majority of DMD patients would read out at 10% of normal dystrophin on IHC. I believe that evidence from Western blot and other experiments discussed above show that protein in the range between undetectable and 10% of normal is likely to be very important for clinical presentation, all other things being equal, i.e., mutation status and non-dystrophin-related factors affecting phenotype.

These findings are germane to the determination of “reasonably likely to predict clinical benefit.” The broad phenotypic distinctions made in the clinic (e.g., DMD vs IMD vs BMD) are different from the prediction of benefit to an individual patient who has a specific baseline dystrophin level and whose mutation and external factors do not change pre- and post-drug. For example, extending ambulation by six months to a year would not normally move a patient from one to another of these categories, but could be very important to quality of life (e.g., as suggested in the Bello study). This is also true for other functional improvements.

For these reasons, incorporating the analysis of dystrophin content discussed above, I conclude that the biochemical data strongly support the idea that low-level increases in dystrophin production are reasonably likely to predict clinical benefit.

Additional support for “reasonably likely” comes from the long-term experience with the drug. The sponsor’s comparison of the experience of the treated cohort to natural history data does not reach the level of substantiation required for traditional approval based on the clinical data. However, it is highly suggestive of improvement in some parameters, in some patients, over natural history. My conclusion is informed by all the caveats expressed in the reviews about the pitfalls of non-randomized comparisons. Given that the two exon 52 deletion patients in the study had fairly good long-term results in terms of rate of disease progression, the question arises as to whether exon 52 is a prognostic factor that could have skewed the results.

Several facts militate against this conclusion. First, one of the exon 52 deletion trial subjects (subject 6) had a fairly low score on the 6MWT at entry and a very low score on the NSAA, compared to other subject around his age. He also was the only subject in the trial noted to be unable to rise without external support at baseline. Additionally, the Italian external cohort had exon 52 deletion representation.

Questions have been raised about the correlation of dystrophin levels from Western blot with clinical outcomes. The 6 Minute Walk Test does not show a strong correlation. I evaluated the NSAA in children who could still walk (because the NSAA primarily scores activities related to walking) and who also had a dystrophin result at 180 weeks (Table 1). I did this because the NSAA includes multiple measures and therefore might have some noise averaged out. I looked at the absolute decline in NSAA in patients since study initiation, and did not correct for the initial time some patients spent on placebo. I only evaluated patients who were ambulatory. There was a positive (inverse) correlation between dystrophin by Western blot and rate of decline in NSAA score, . (Figure 1) This adds additional support to the idea that dystrophin production is “reasonably likely to predict clinical benefit.” In totality, I find that the comparative disease course data provide additional support for the use of the surrogate endpoint of an increase in dystrophin expression as “reasonably likely to predict clinical benefit.”

Therefore, both the biochemical data and the clinical data lead me to conclude that an “increase in dystrophin production” is reasonably likely to predict clinical benefit in DMD.

CONFIRMATORY TRIALS

The sponsor is currently conducting a nonrandomized, concurrently controlled trial in patients with mutations amenable to exon 51 skipping compared to untreated DMD patients with other exon deletions. Because of the relatively low level of protein induced, additional doses should be aggressively pursued and, if successful, a dose-comparison trial could be confirmatory. The sponsor has also planned to initiate a randomized trial with a related compound in other exons. The clinical results from these trials can inform the predictive value of the surrogate endpoint.

EXPLORATION OF ADDITIONAL DOSES, REGIMENS, AND DRUG-MUTATION INTERACTION

The dystrophin levels achieved in this development program are well below those initially hoped for. I agree with Dr. Farkas and other reviewers that the sponsor should aggressively explore higher doses or more frequent administration of eteplirsen. It appears that this is possible given the toxicology data and the clinical safety profile observed to date.

Because patients in the Sarepta 180 week cohort had a range of deletions in the dystrophin gene, variability in the pharmacodynamic response among deletions is of great interest. The two patients with over 2% dystrophin in the 180 week Western blot both had exon 52 deletions. These patients also fared fairly well, clinically. This raises the question of whether patients with this exon deletion naturally produce more dystrophin. One of these subjects had a baseline sample available. It was found to be below the limit of quantitation. There was an exon 52 subject included in the added baseline controls. This subject's assay had replicate results of 0.3% and below the limit of quantification, respectively, as discussed above. This suggests that baseline dystrophin levels are not higher in exon 52 deletion subjects and that there may be a drug-deletion interaction, wherein subjects with this deletion may have a more robust pharmacodynamic response to the drug. There were a number of apparent non-responders to the drug. It will be important to find out if this is mutation specific. It is likely that more detailed knowledge about each patient's specific mutation will have to be generated to study this in detail.

COMMENTS ON THE DEVELOPMENT PROGRAM AND REVIEW

The development program for eteplirsen was seriously deficient in a number of respects that may have led to delay in broad access and certainly led to difficulties in regulatory review. In my assessment, the most egregious flaw was the lack of robust and high-quality assays early in the development program. Inaccurate conclusions from the assays used led to a flawed development program. Additionally, the entire drug development field must recognize that there is no such thing as an "exploratory study" for a serious, life-threatening illness without therapeutic options. Randomization should be performed very early in the development program, and open-label studies should be avoided. When possible, seamless adaptive dose-finding and early efficacy studies should be carried out with the goal of most efficiently generating the data needed to demonstrate safety and effectiveness.

The flaws in the eteplirsen development program led to severe challenges in regulatory review. 21 CFR 312.80, concerning drugs intended to treat life-threatening or severely-debilitating illness, states that FDA has determined "that it is appropriate to exercise the broadest flexibility in applying the statutory standards, while preserving appropriate guarantees for safety and effectiveness...Physicians and patients are generally willing to accept greater risks or side effects from products that treat life-threatening and severely-debilitating illnesses than they would accept from products that treat less serious illnesses." I note that the acceptable risks include greater uncertainty about the effects of the drug. The Peripheral and Central Nervous System Drugs Advisory Committee met on this application on April 25, 2016. There was a split vote (7 against, 6 for) on the question of accelerated approval for this drug, reflecting the greater than usual uncertainty about the application. This vote was taken before the additional data on protein expression were submitted.

To conclude, the studies used in this analysis to support the effect of eteplirsen on dystrophin were adequate and well-controlled as specified in 314.126. In addition, the surrogate of increased dystrophin production is reasonably likely to predict clinical benefit. Given the deficiencies that have been identified in the development program, my conclusion to rely on the surrogate endpoint described above represents the greatest flexibility possible for FDA while remaining within its statutory framework. In this case, the flexibility is warranted because of several specific factors, including: the life-threatening nature of the disease; the lack of available therapy; the fact that the intended population is a small subset of an already rare disease; and the fact that this is a fatal disease in children. Of note, the therapy has been relatively safe in the clinic, although intravenous administration always carries risk. In addition, adequate confirmatory studies are underway and planned and are capable of further refining our understanding of the biomarker and providing evidence about the nature of the clinical benefit. The approval does not create any risk of compromising the confirmatory trials because of their nature. Therefore, I find that the probable benefits outweigh the foreseeable risks and that this application should be approved under 21 CFR 314.510.

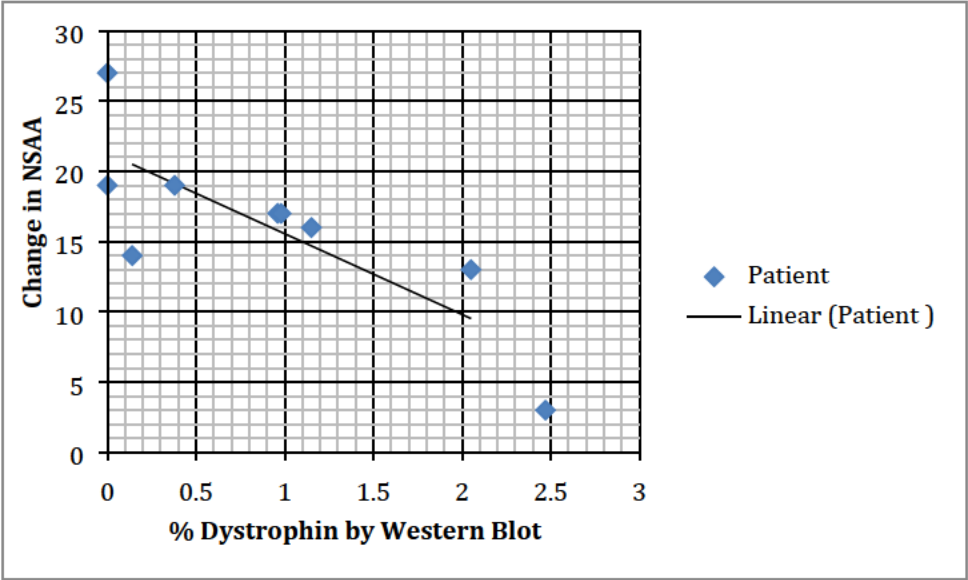
Table 1 Patient Data on Change from Baseline in 6MWT and NSAA

180 Weeks
Change from
Baseline

Subject	Baseline WB	180 Week WB	Fiber Intensity	PDPF	Δ 6MW	Δ NSAA
002	N/A	0.14	MD	4.54	-67	-14
003	N/A	0	7	1.42	-174	-27
004	N/A	0.96	28	29	-168	-17
005	0	N/A	N/A	N/A	-231	-19
006	N/A	2.47	29	21	-23	-3
007	N/A	0	12	7	-197	-19
008	N/A	0.98	26	12	-291	-17
009	N/A	0.52	23	22	0	0
010	N/A	1.62	21	24	0	0
012	N/A	0.38	26	33	-114	-19
013	0	1.15	32	19	-170	-16
015	0	2.05	30	18	-1	-13

Data from Sarepta Therapeutics, Inc. PCNSD Advisory Committee Briefing Document, Appendix 5, p. 149 (6MW and NSAAO Appendix 11, p. 155, (Percent Positive Dystrophin Fibers (PPDF), Appendix 12 p. 156 (fiber intensity) 14, p. 159. (Western blot),

Figure 1. Decline in NSAA by % Dystrophin on Western blot



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

/s/

JANET WOODCOCK
07/14/2016

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 Laurensen, 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) (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.

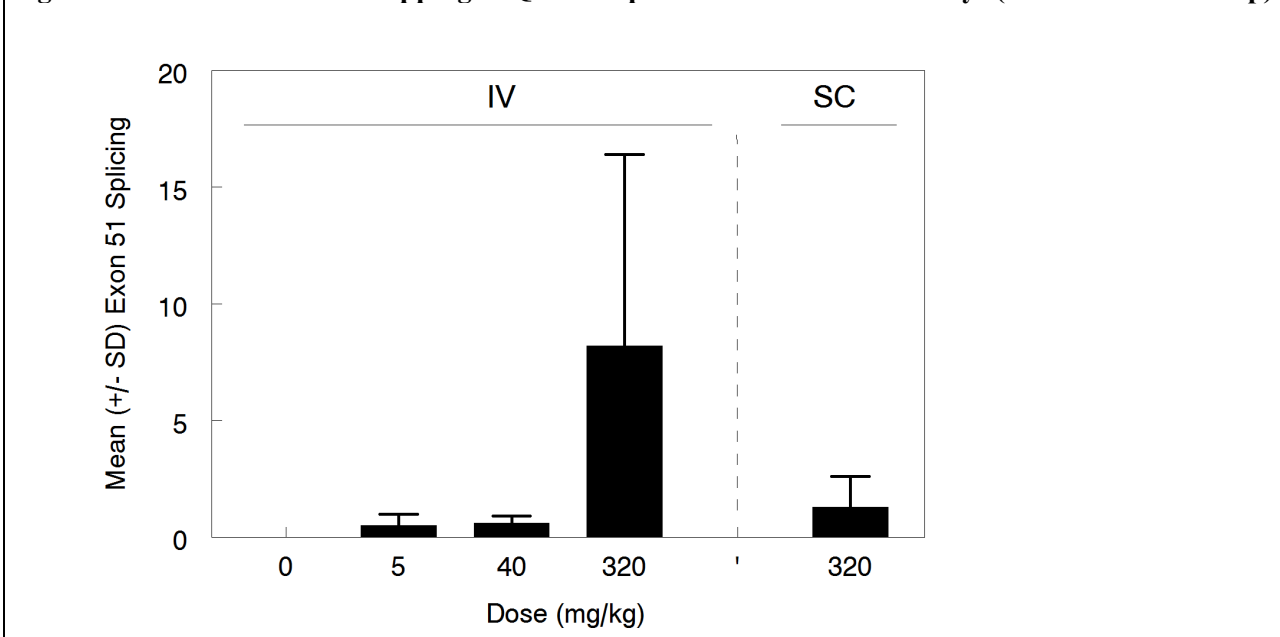
The applicant presented data on the exon skipping activity of eteplirsen in cynomolgus monkeys (“Exon skipping activity of AVI-4658 in cynomolgus monkey tissue samples from ^{(b) (4)} study ^{(b) (4)},” applicant study 4658-ssa-005). Samples of quadriceps muscle, heart, and diaphragm tissues were collected 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.

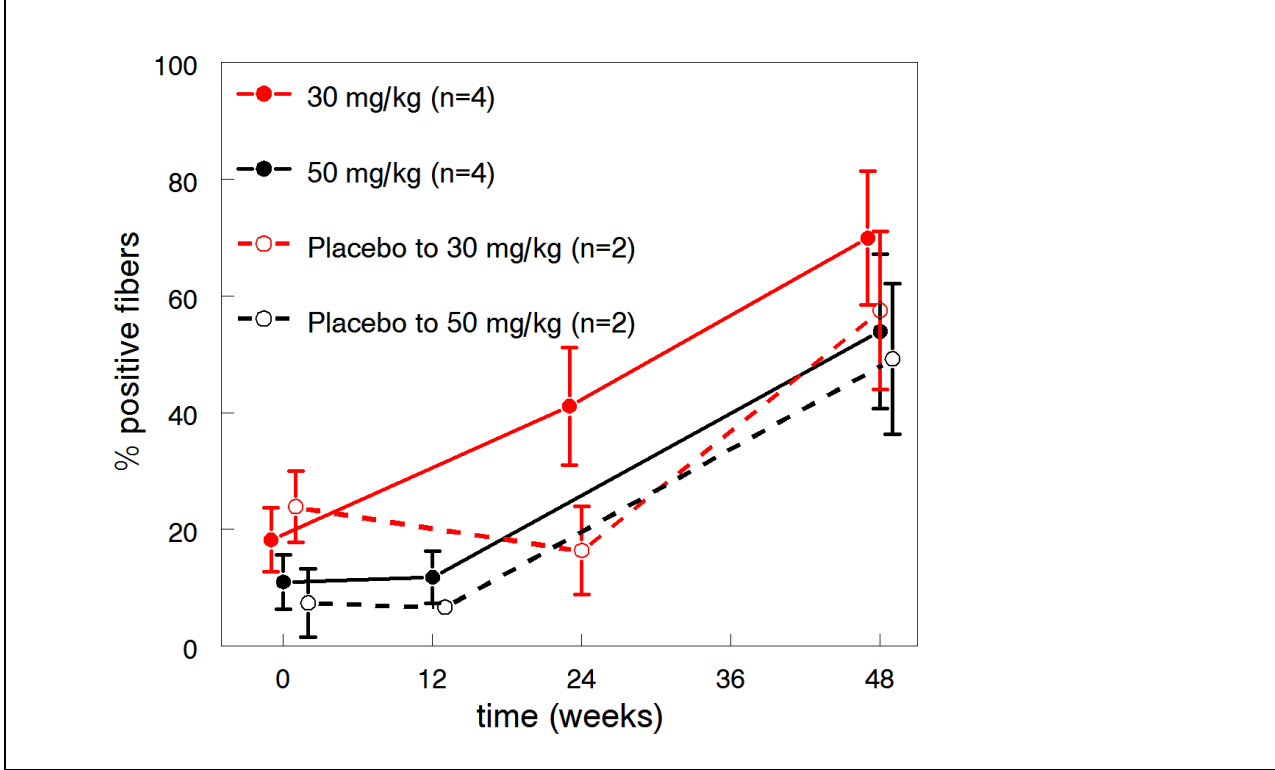
Publication of this paper was followed by a Sarepta press release,³ which also claimed a remarkable treatment effect from eteplirsen and raised wildly unrealistic expectations in the DMD community. It was these perceptions and expectations that led the applicant to declare that a placebo-controlled study 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



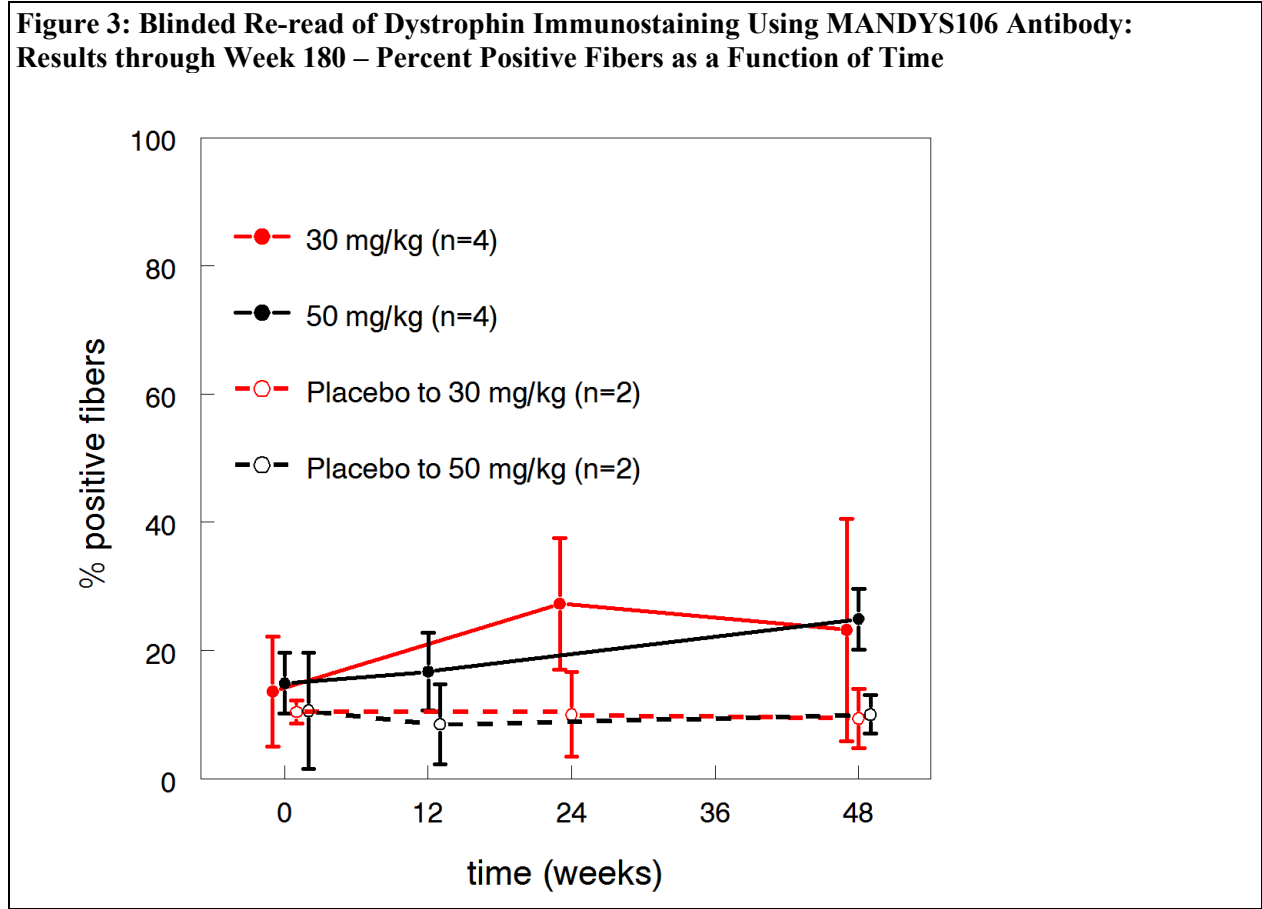
I was part of an inspection team that conducted (May 29 and 30, 2014) a site visit to Nationwide Children’s Hospital in Columbus, OH, where Study 201 was conducted. We found the analytical procedures to be typical of an academic research center, seemingly appropriate for what was simply an exploratory 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.

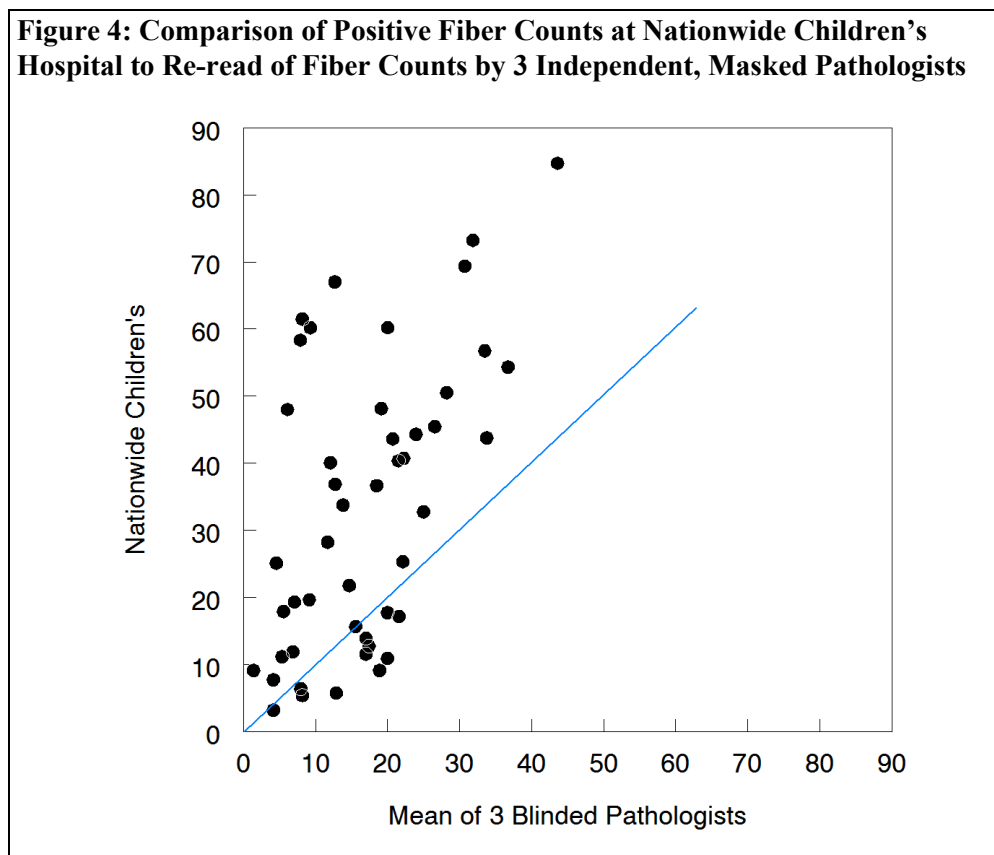
Having uncovered numerous technical and operational shortcomings in Columbus, our team worked collaboratively with the applicant to develop improved methods for a reassessment of the stored images. We suggested a re-read of all images by 3 independent masked readers, such that blinding could be assured 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.

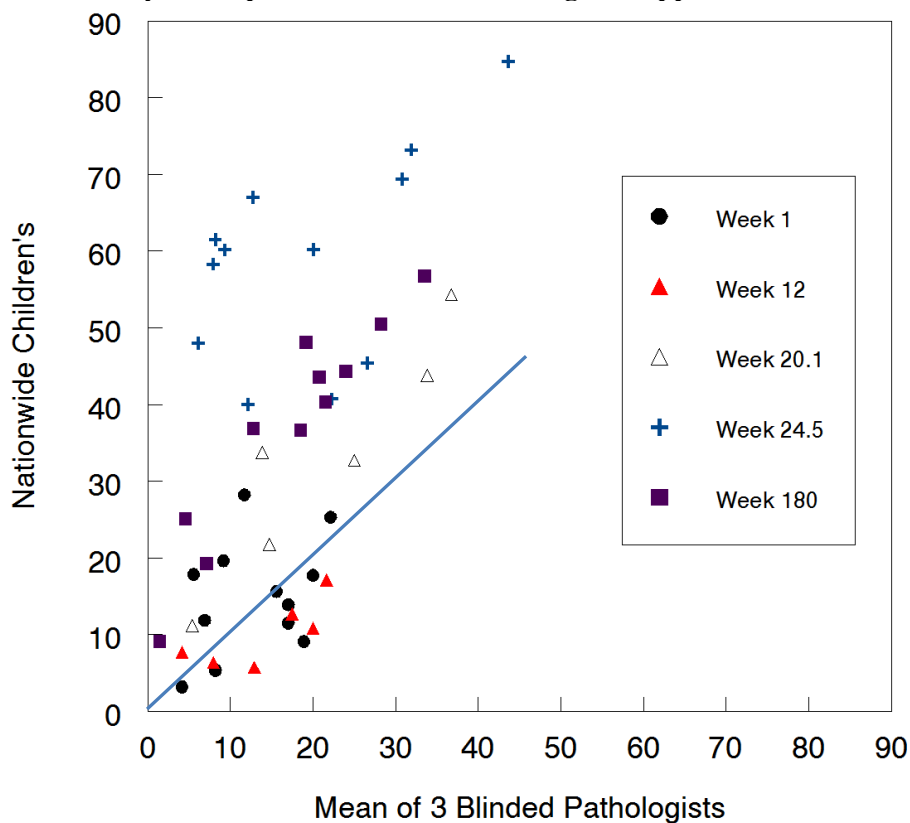


One might reasonably ask why the original readings were not reproduced by a blinded re-read. Figure 5 shows the same scatterplot between readings by Nationwide Children’s Hospital and the 3 blinded pathologists. In this display, however, readings from samples obtained at the disparate time points are 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 ($1.1 \pm 1.3\%$).

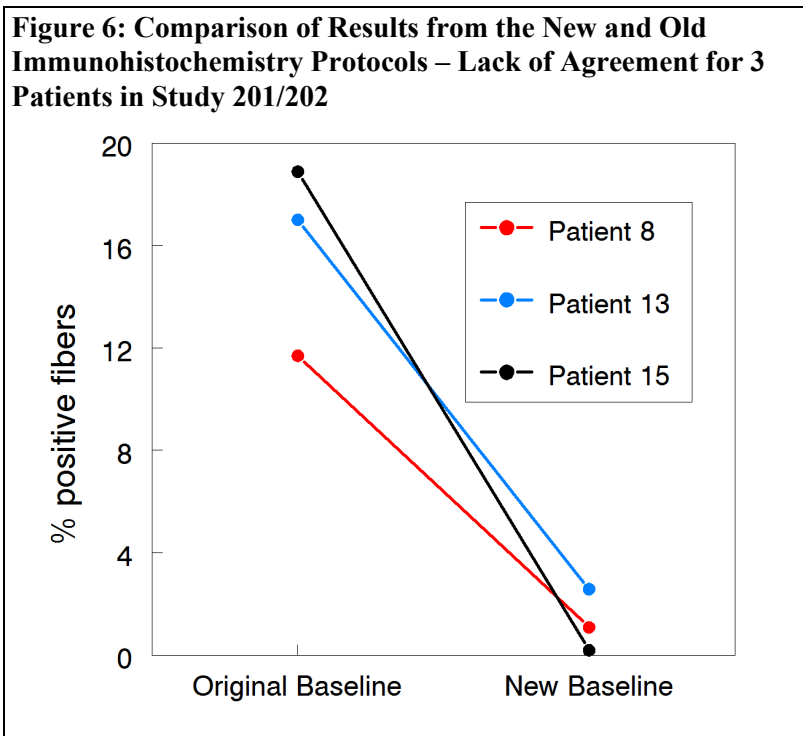
As noted above, there were 3 patients in Study 201/202 with adequate archived tissue for separate immunohistochemistry analyses using both the old and new methods. Figure 6 shows how the two methods compare. These are essentially replicate analyses of a single tissue sample using the two methods.

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.

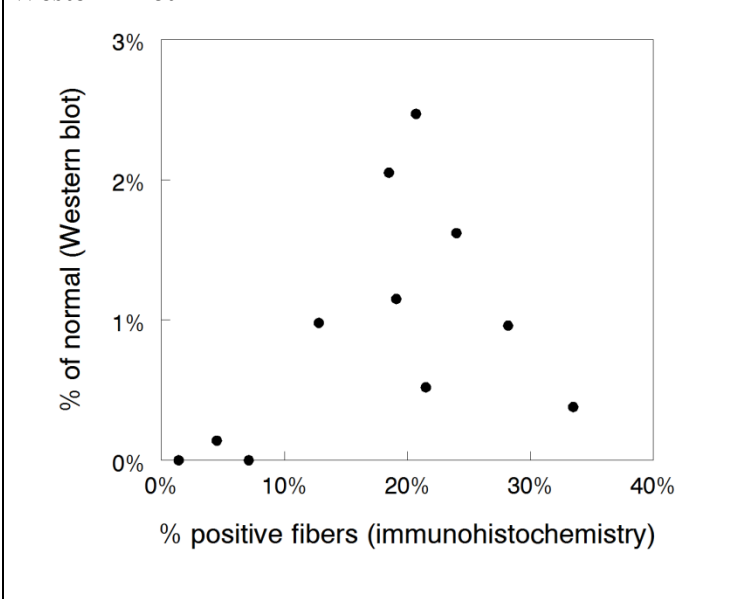
The age of these 13 patients ranged from 7 to 13 years. Paired pre- and post-treatment samples were run in side-by-side lanes on the gels, and each gel was run in duplicate. A muscle sample from a healthy 14 year-old boy with no pathologic diagnosis served as the reference sample; values from the DMD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

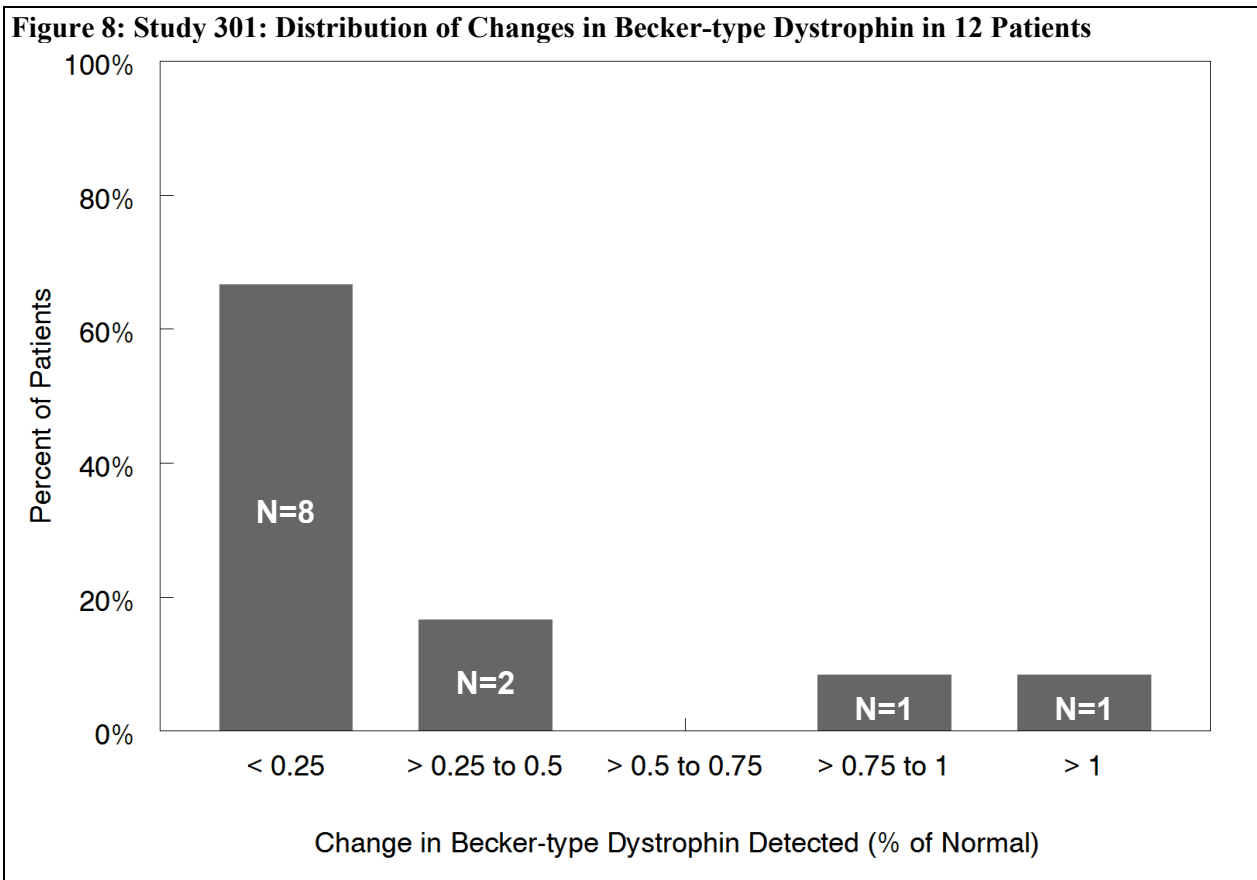
Reporting values below the limit of quantification as 0, pre- and post-treatment values are $0.06\% \pm 0.14\%$ and $0.38\% \pm 0.50\%$, respectively (mean \pm standard deviation), $p < 0.05$. For the 'as reported' analysis, pre- and post-treatment values are $0.16\% \pm 0.12\%$ and $0.44\% \pm 0.43\%$, respectively, $p < 0.05$. Reporting all values below the limit of quantification as 0.24%, pre- and post-treatment values are $0.26\% \pm 0.05\%$ and $0.48\% \pm 0.41\%$, 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	0.08		
	Week 48	pass	0.5	0.37			Week 48	pass	0.31	0.47	
		pass	0.24					pass	0.63		
4	Baseline	pass	0.04	0.04	0.06	11	Baseline	pass	0.01	0.02	0.07
		fail	0.06					pass	0.09		
	Week 48	pass	0.1	0.1			Week 48	fail	0.01	0.09	
		fail	0.19					pass	0.01		
5	Baseline	fail	0.1	0.17	0.85	12	Baseline	pass	0.02	0.18	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.

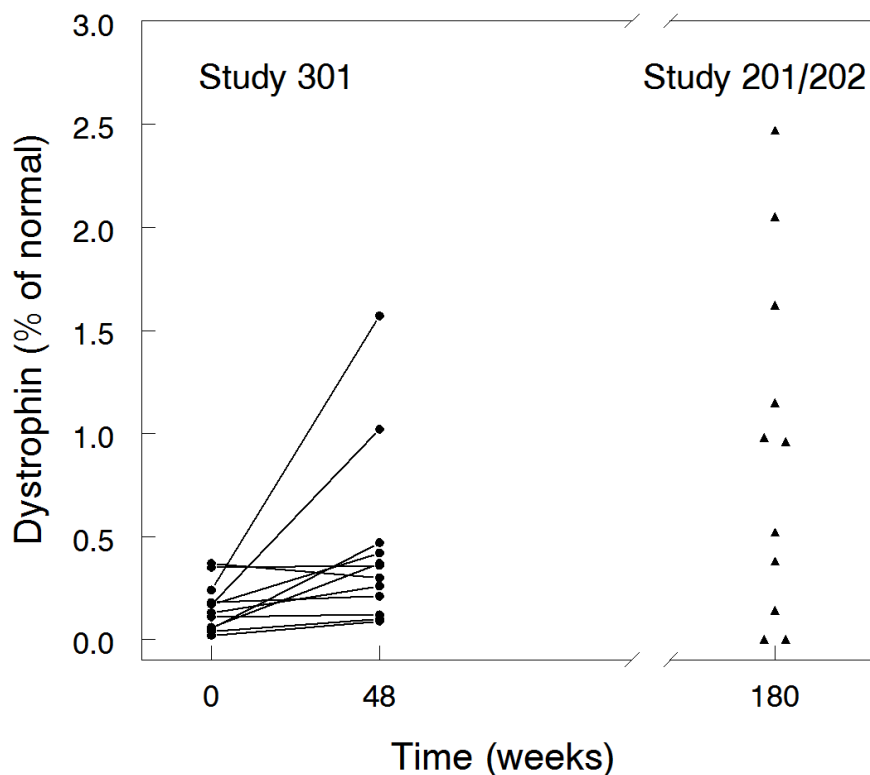
The problem with a responder analysis is that there is no rational basis upon which to define a threshold for a ‘response.’ Various cut-points could be selected, but their selection would be arbitrary, and the particular threshold chosen would have a major influence on the effect size.

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.

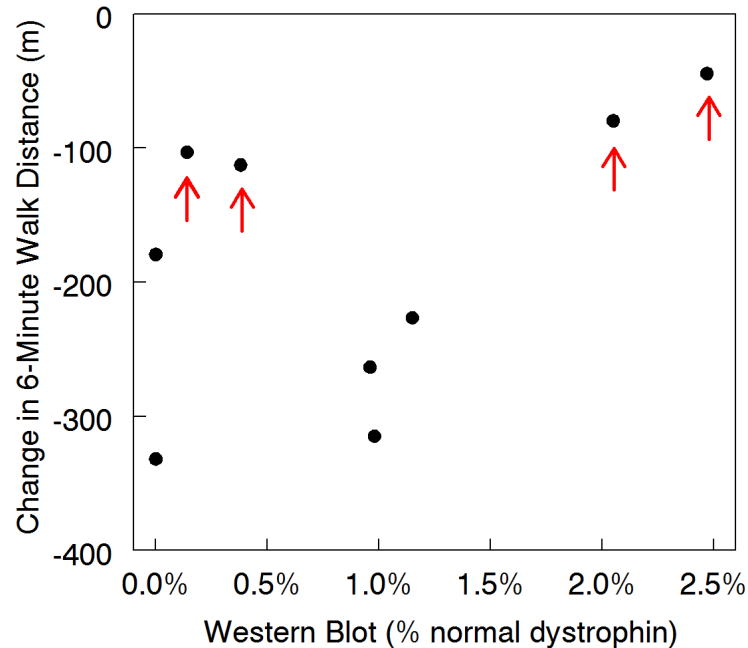
Do the clinical data bolster the question of “reasonably likely?”

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.

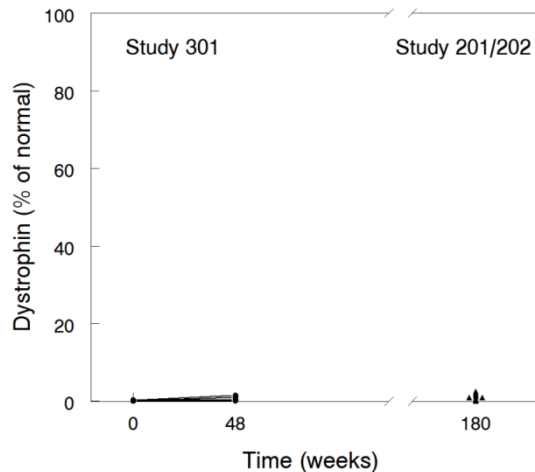
Figure 10: Study 201/202 – Lack of Correlation between Quantity of Dystrophin Detected and Preservation of Physical Function (6-Minute Walk Distance)



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.’

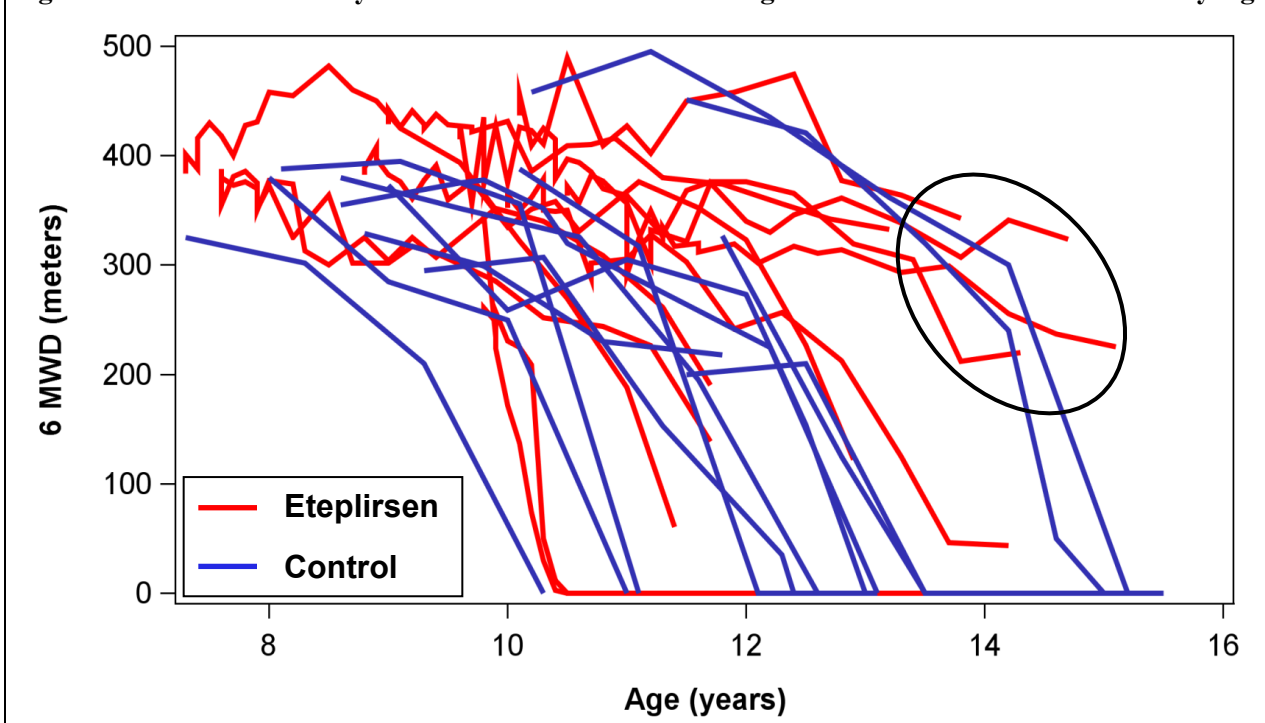
Study 202 was continued, therefore, with patients continuing to receive either 30 or 50 mg/kg/week eteplirsen. Numerous comparisons of physical function were planned between these 12 patients and the 13 patients in the external control group. Measures included 6-minute walk, rise time, timed 10-meter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



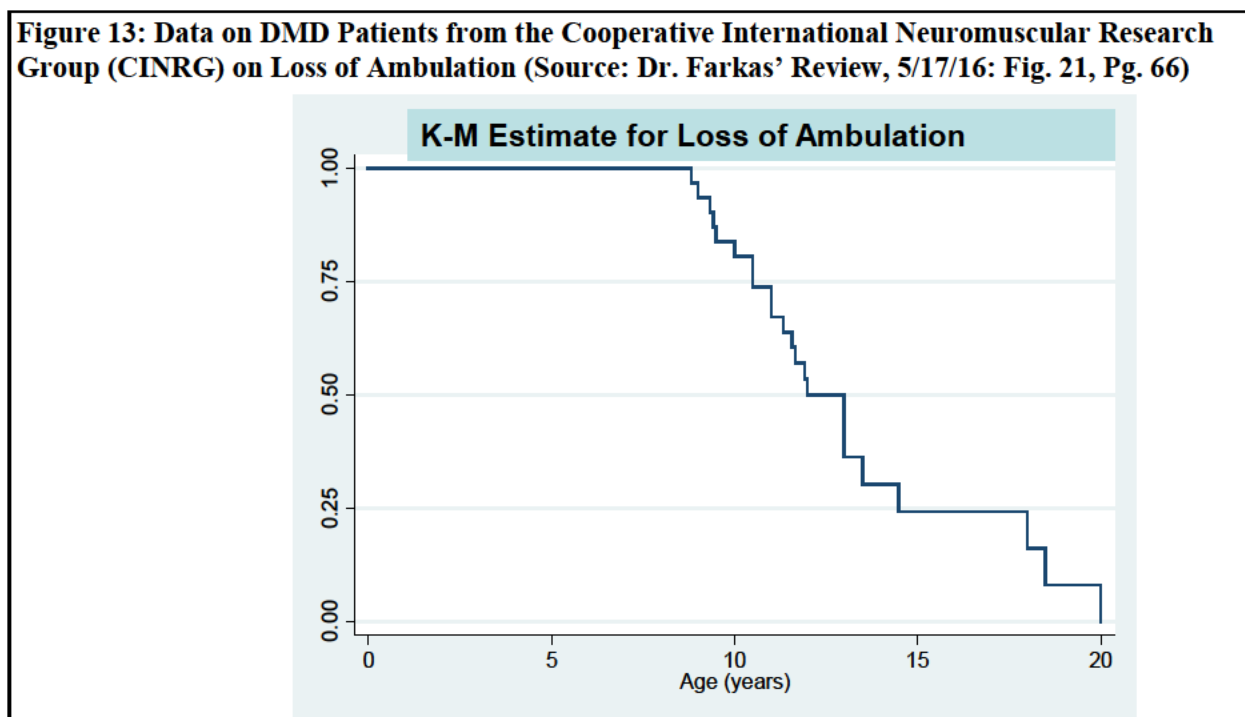
by simple visual inspection, the two groups show little difference in performance.

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.

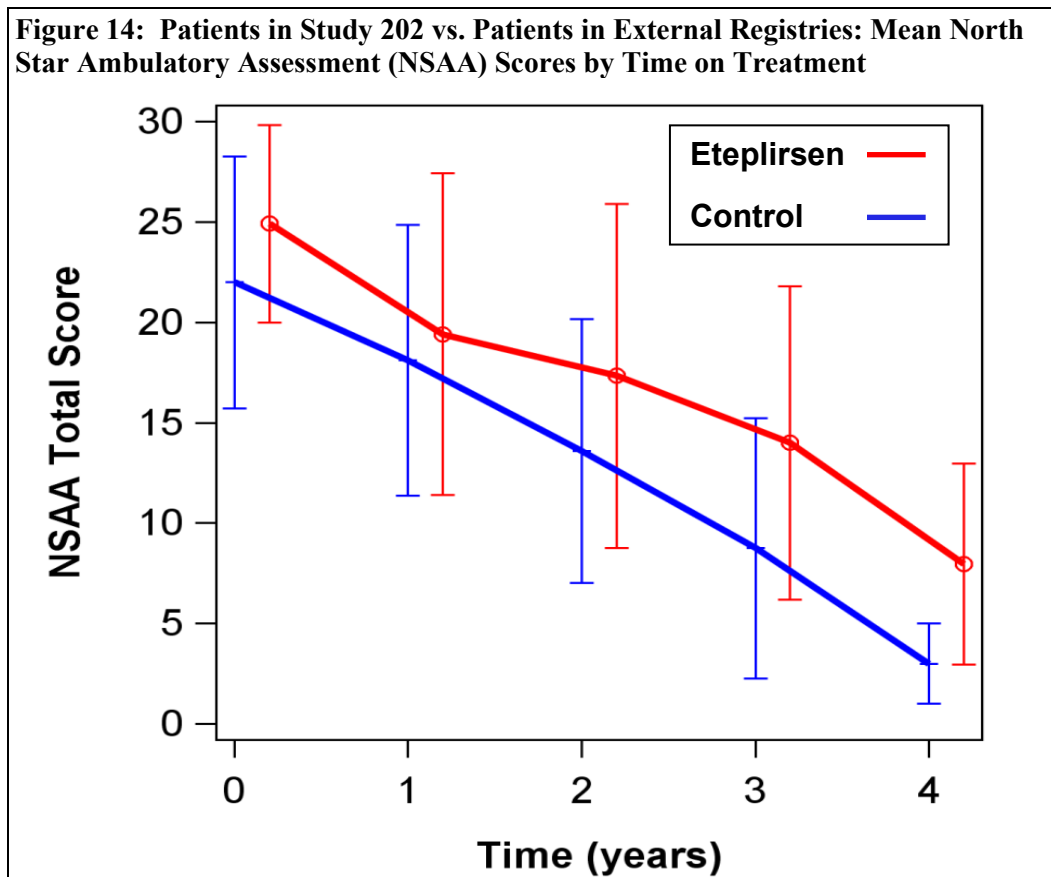
Data from the Cooperative International Neuromuscular Research Group (CINRG) provide an additional source of information on the natural history of patients with DMD. Figure 13 is a Kaplan-Meier (K-M) survival curve from CINRG showing time-to-loss of ambulation. Of note, 25% of patients remain ambulatory 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.

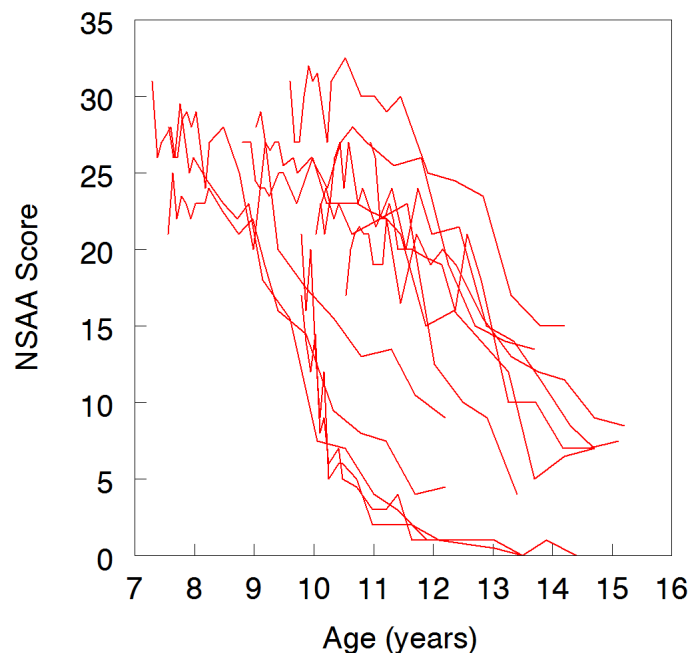
Importantly however, despite the claims of improvement made at the microphone at the Advisory Committee meeting, the review team did not find any patients in Study 201/202 with consistent improvement in physical performance as assessed by formal testing (6-minute walk, rise time, NSAA, 10-meter run). 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