3

A community-engaged approach to quantifying caregiver preferences for the benefits and risks of emerging therapies for Duchenne muscular dystrophy

A Community-Engaged Approach to Quantifying Caregiver Preferences for the Benefits and Risks of Emerging Therapies for Duchenne Muscular Dystrophy

Holly L. Peay, MS[1]; Ilene Hollin, MPH[2]; Ryan Fischer, BA[1]; and John F.P. Bridges, PhD[2]

[1]Parent Project Muscular Dystrophy, Hackensack, New Jersey; and [2]Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, Maryland

ABSTRACT

Background: There is growing agreement that regulators performing benefit–risk evaluations should take patients' and caregivers' preferences into consideration. The Patient-Focused Drug Development Initiative at the US Food and Drug Administration offers patients and caregivers an enhanced opportunity to contribute to regulatory processes by offering direct testimonials. This process may be advanced by providing scientific evidence regarding treatment preferences through engagement of a broad community of patients and caregivers.

Objective: In this article, we demonstrate a community-engaged approach to measure caregiver preferences for potential benefits and risks of emerging therapies for Duchenne muscular dystrophy (DMD).

Methods: An advocacy oversight team led the community-engaged study. Caregivers' treatment preferences were measured by using best–worst scaling (BWS). Six relevant and understandable attributes describing potential benefits and risks of emerging DMD therapies were identified through engagement with advocates (n = 5), clinicians (n = 9), drug developers from pharmaceutical companies and academic centers (n = 11), and other stakeholders (n = 5). The attributes, each defined across 3 levels, included muscle function, life span, knowledge about the drug, nausea, risk of bleeds, and risk of arrhythmia. Cognitive interviewing with caregivers (n = 7) was used to refine terminology and assess acceptability of the BWS instrument. The study was implemented through an online survey of DMD caregivers, who were recruited in the United States through an advocacy group and snowball sampling. Caregivers were presented with 18 treatment profiles, identified via a main-effect orthogonal experimental design, in which the dependent variable was the respondents' judgment as to the best and worst feature in each profile. Preference weights were estimated by calculating the relative number of times a feature was chosen as best and as worst, which were then used to estimate relative attribute importance.

Results: A total of 119 DMD caregivers completed the BWS instrument; they were predominately biological mothers (67.2%), married (89.9%), and white (91.6%). Treatment effect on muscle function was the most important among experimental attributes (28.7%), followed by risk of heart arrhythmia (22.4%) and risk of bleeding (21.2%). Having additional postapproval data was relatively the least important attribute (2.3%).

Conclusions: We present a model process for advocacy organizations aiming to promote patient-centered drug development. The community-engaged approach was successfully used to develop and implement a survey to measure caregiver preferences. Caregivers were willing to accept a serious risk when balanced with a noncurative treatment, even absent improvement in life span. These preferences should inform the Food and Drug Administration's benefit–risk assessment of emerging DMD therapies. This study highlights the synergistic integration of traditional advocacy methods and scientific approach to quantify benefit–risk preferences. (*Clin Ther.* 2014;36:624–637)

Key words: benefit–risk assessment, caregiver, choice behavior, Duchenne muscular dystrophy, patient preferences.

Accepted for publication April 9, 2014.
http://dx.doi.org/10.1016/j.clinthera.2014.04.011
0149-2918/$ - see front matter

INTRODUCTION

Duchenne muscular dystrophy (DMD) is a rare, life-threatening, inherited neuromuscular disorder that occurs in male subjects with an incidence of 1.3 to 2.9 per 10,000.[1,2] Diagnosis usually occurs around age 5 years, when differences in motor function become apparent, but symptoms may appear as early as infancy.[3,4] The condition is associated with significant care-related[5,6] and financial burden.[7–9] Affected individuals have progressive muscular weakness, loss of ambulation that typically occurs in the teen years, and premature death.[3] The mean age of death is in the 20s and is commonly caused by respiratory failure or cardiac disease.[3,4]

Currently, there are no therapies approved by the US Food and Drug Administration (FDA) for DMD. The standard-of-care treatment is the off-label use of corticosteroids, which have been shown to stabilize muscle strength, delay loss of ambulation by 2 to 5 years, improve cardiopulmonary function, and enhance quality of life.[3,10–12] Several potential therapies are under clinical trial that target a variety of primary and secondary effect pathways.[13]

Similar to other conditions (including other rare diseases[14] and early-on in the HIV epidemic),[15] patients and caregivers managing DMD seek to accelerate approvals for drugs that may save lives.[16] In the context of serious, rare disorders with limited treatment options, patients and patient advocates want regulators to be more permissive.[17] Drugs under trial for DMD represent a significant opportunity for families to intervene, and in public forums, some parent advocates have demanded access to drugs, even absent conclusive data on efficacy and safety.[18] DMD provides an appealing model for assessing influences on treatment decision making for serious, progressive disorders. The natural history may lead to high-pressure decisions regarding the use of novel therapies. The pediatric onset provides additional complexity, as the majority of treatment decisions are made by parents/guardians who are also the primary caregivers. These decisions may later be re-evaluated by adolescent and adult DMD patients who could have different treatment preferences.[19]

There is growing agreement that regulatory benefit–risk evaluations should be informed by the perspectives of patients and caregivers who will ultimately make treatment decisions and bear the associated risks.[20] To that end, in 2012, the FDA was congressionally mandated to engage patients to understand the impact of disease though the Patient-Focused Drug Development Initiative.[21] Although this program offers patients and caregivers an unprecedented opportunity to contribute to the regulatory process, the program is limited in scope and approach, with initially only 20 disease areas being targeted for public comment involving direct engagement with patients and caregivers.[22] Advocates who do not represent 1 of the 20 chosen disorders are left with little guidance about how to provide input that is acceptable and useful to the FDA. Existing models for FDA engagement are largely limited to providing testimonial. Although such direct engagement is a primary strength and a mainstay of advocacy organizations' efforts to inform decision makers about their community's needs and perspectives,[23] there are limitations to focusing only on patient and caregiver testimonials, such as questions about how well those providing testimonial represent the views of the entire patient population.[20,24,25] Increasingly, decision makers are being asked to consider alternative methods to quantify treatment preferences and risk tolerance that take into account the views of large groups of stakeholders.[26] Draft FDA guidance has indicated a willingness to incorporate such evidence into the regulatory process.[27] Quantitative preference elicitation methods allow stakeholders to introduce formal evidence-based decision making into the regulatory process and have been used to explore decision making and preferences among a variety of patient populations.[20]

The purpose of the present study was to demonstrate a process by which a patient advocacy organization might develop scientific evidence on treatment preferences. We aimed to model a replicable, community-engaged approach to exploring preferences in a large sample of decision makers. Specifically, our goal was to explore caregiver preferences for emerging treatments for DMD. This study is not only informative to those seeking to understand the treatment preferences and risk tolerance of DMD caregivers, but it serves to highlight principles of patient-centered outcomes research[28] by illustrating how an advocacy organization can take leadership in generating policy-relevant evidence.

MATERIALS AND METHODS

Parent Project Muscular Dystrophy (PPMD), an advocacy organization focused on finding a cure for

DMD, led the study. The advocacy oversight team comprising PPMD staff members (a clinician, a scientist experienced in drug development, and 2 caregivers of individuals with DMD) collaborated with the research team to design and implement the study. The oversight team made study decisions through a consensus process. Consistent with the preferences of PPMD, the authoring research team was a smaller team of PPMD staff and academic collaborators.

The teams began by defining a research question about treatment preferences based on the stated needs of the Duchenne community; they then identified the study population (caregivers of individuals with DMD) and a recruitment strategy. The teams choose a stated preference method (described under the heading "Methods") that fit the study needs. In the development of the treatment experiment (described under the heading "Identifying Attributes and Levels"), the team used a community-engaged approach involving multidisciplinary stakeholder informants. The survey was piloted by a small group of end-users (described under the heading "Survey Pilot: Cognitive Interview"), and it was modified based on their input. Preliminary and final analyses were reported to regulators, industry, and the Duchenne community in an accessible and timely manner.

Methods

Methods to measure the preferences of patients, caregivers, and other stakeholders are now well established[29,30] and are increasingly being applied to study benefit–risk preferences.[20] Although good research practices have been created to aid in the development of stated preference applications in medicine,[31] approaches such as conjoint analysis and discrete choice experiments remain complex. They require qualitative skills to appropriately identify attributes and levels and develop supporting survey text,[32] as well as quantitative skills to design the experiment[33] and analyze results.[34]

The study used best–worst scaling (BWS) case 2, an emerging stated-preference method that can be used to scientifically assess preferences.[35] Referred to as the "profile case," BWS case 2 presents profiles one at a time to elicit preferences; Flynn et al[36] fully described the method and provide use guidance. BWS has been recognized as an approach that is easier to design and analyze than conjoint analysis and discrete choice; however, this method is relatively novel in the context of measuring benefit–risk preferences.[20] The study reported here presents a novel use of the BWS case 2.[37,38] BWS is thought to be less cognitively demanding on participants than discrete choice or conjoint experiments.[36] In addition, relevant to our aim of demonstrating a replicable community-engaged model, BWS benefits from a straightforward analytic approach, the results of which are consistent with more complex approaches[39,40] that may be unfamiliar to many researchers in the clinical domain. To guide the development, implementation, and analysis of our preference elicitation instrument, we used the standards outlined in the International Society for Pharmacoeconomics and Outcomes Research checklist for conjoint analysis[27] and specific guidance on the use of BWS.[34]

Identifying Attributes and Levels

BWS case 2 experiments use attributes (representing topic areas) and levels (representing attribute variables, such as amount of or impact on the attributes); Figure 1 presents an example of a BWS case 2 task. Identification of relevant and comprehensible attributes and levels is required for a meaningful study outcome. We used a stakeholder-informed approach to identify attributes and levels that were clinically relevant, meaningful, and understandable to caregivers. The development of attributes and levels was informed by PPMD's 20 years of experience with patients and families; extensive history consulting on, reviewing, and funding clinical research; and an ongoing interview study of clinical trial experiences.[41]

PPMD identified and invited stakeholder informants to participate (October–December 2012) through an existing advocacy-facilitated industry roundtable, PPMD's grassroots family networks, PPMD's clinician database, or after self-nomination following community notification of the program launch. Stakeholder informants, including patient/disease advocates (n = 5), clinicians (n = 9), drug developers from pharmaceutical companies and academic centers (n = 11), and other stakeholders (n = 5), participated in group or individual sessions. Attributes and levels for the emerging therapies were proposed and refined through iterative rounds of stakeholder engagement. Industry informants were important to successfully identify appropriate attributes because of their ability to forecast benefits and risks of premarket drugs. The

Choose the best thing in this treatment by clicking the circle under "best" and choose the worst thing by clicking the circle under "worst." You have to choose a best thing and worst thing to move on. Remember that a computer chose the combinations to make the experiment work, and some of them seem bad. Even so, please pick the best and worst thing.

Best	Treatment	Worst
○	Slows the progression of weakness	○
○	2-year gain in expected life span	○
○	1-year of postapproval drug information available	○
○	Causes loss of appetite	○
○	Increased risk of bleeding gums and increased bruising	○
○	Increased risk of harmless heart arrhythmia	○

Figure 1. Example of best–worst scaling task.

oversight committee incorporated their informative input while protecting against potential bias.

This approach yielded >20 potential benefit, adverse effect, and risk attributes. Items were grouped under themes and refined. Several attributes were rejected by the oversight team for relevance, similarity to other items, or concerns about the ability of the target population to understand the attribute. Examples of attributes that were not chosen are the ability to participate in day-to-day family activities and risks to renal and hepatic function. The participation benefit was considered to be less concrete and treatment-associated than muscle function, and thus was not chosen. The clinical implications of renal and hepatic damage were difficult to describe in a brief and accessible format. In addition, we chose not to use quantitative risks in this survey because the average US adult has only a basic level of quantitative literacy.[42]

Through a consensus process, the advocacy oversight team ultimately selected 6 treatment attributes with 3 levels each. The attributes and associated levels were chosen to be reasonable based on drugs under trial, with the notable exception that the highest risk levels represent considerably more risk than has been associated with drugs under trial, to date. The proposed attribute list was again shared with stakeholder informants. Based on their input, the final items (effect on muscle function, life span, knowledge about the drug, nausea, risk of bleeding, and risk of heart arrhythmia) were chosen by the oversight group and study team.

The BWS experiment attributes and levels are presented in Table I. The functional benefits chosen were "stops the progression of weakness" and "slows the progression of weakness" because drugs under trial are unlikely to result in a cure or significant improvement in strength for patients with DMD. The life span attribute was presented independently of the weakness attribute because drugs that affect skeletal muscle may not improve cardiac outcomes in DMD[43] and thus may not improve longevity. Caregiver participants were prompted to separate muscle function from life span by use of this cardiac example and a sample task. Given that the quality of evidence and associated uncertainty may affect preferences, we included an attribute relating to knowledge about the drug, described as the number of years of postapproval data. Nausea represents a realistic, easily understood adverse effect that may result in increasing burden as patients lose mobility. The choice of bleeding was prompted by a Phase II DMD trial that was terminated in 2011 (unpublished data). Arrhythmia was chosen as an attribute that is salient to caregivers because it is part of the DMD natural history.[44]

Experimental Design

Following good research practices, we developed the BWS case 2 experiment to accommodate the 6 attributes with 3 levels each.[31] We applied a 3^6 main-effects orthogonal design, identified from the SAS database of orthogonal arrays.[45] This array consisted of 18 full-profile combinations of the attributes and levels, the minimum such number necessary to ensure no structural relationships (ie, correlations) between the attributes.[33]

As illustrated in Figure 1, in each BWS task, caregivers were presented with one of the treatment profiles and asked to judge which aspects they thought were the best and the worst. Before completing the tasks, caregivers were presented with a detailed description of all the attributes and levels to be considered in the task

Table I. Attribute, levels, and descriptions resulting from stakeholder engagement.

Attributes and Attribute Lvels	Additional Description/Explanation in the Survey
Effect on muscle function	"How the treatment affects muscle function"
i Stops the progression of weakness	"Most people who take this treatment don't get any weaker over time.
ii Slows the progression of weakness	"Most people who take this treatment continue to get weaker over time, but more slowly than they would without treatment."
iii Does not change progression of weakness	
Lifespan	"By this we mean how many extra years of life are expected because of the treatment."
i 5 year gain in expected lifespan	"5 extra years of life"
ii 2 year gain in expected lifespan	"2 extra years of life"
iii No extra gain to expected lifespan	"0 extra years, meaning that the treatment may not change the person's lifespan at all."
Knowledge about the drug	"Important information about treatments comes after FDA approval, from tracking people who take the treatment over time. Tracking helps us better understand benefits, risks, potential drug interactions, and how the treatment affects people of different ages and stages of progression. Imagine that everyone who takes the treatment is tracked."
i 2 years of post-approval drug information available	"The treatment has been on the market for 2 years and we have data from 900 people with Duchenne."
ii 1 year of post-approval drug information available	"The treatment has been on the market for one year and we have data from 200 people with Duchenne."
iii No post-approval drug information available	"The treatment has just been approved and no post-approval data is available."
Nausea	
i No increased chance of nausea	
ii Causes loss of appetite	"A person taking the treatment loses his/her appetite"
iii Causes loss of appetite with occasional vomiting	"A person taking the treatment loses his/her appetite and has occasional vomiting"
Risk of bleeds	"Risk of bleeding"
i No increased risk of bleeds	
ii Increased risk of bleeding gums and increased bruising	"Bleeding gums and increased bruising, without increased risk of more dangerous bleeding"
iii Increased risk of hemorrhagic stroke and lifelong disability	"Hemorrhagic (bleeding) stroke, which could lead to lifelong disability in memory and reasoning. People found to have this risk would have to stop taking the treatment."
Risk of heart arrhythmia	"Risk of heart rhythm problems"
i No increased risk of heart arrhythmia	
ii Increased risk of harmless heart arrhythmia	"Occasional, harmless heart arrhythmia"
iii Increased risk of dangerous heart arrhythmia and sudden death	"Dangerous arrhythmia, which could lead to surgery to put in a defibrillator and risk of sudden death. People found to have this risk would have to stop taking the treatment."

(including warm-up questions where appropriate), detailed instructions, and an explained example task. Furthermore, we confirmed that these "treatments do not currently exist" and that we were "interested in knowing what [the caregiver] would choose if they did." As a matter of context, we informed the caregivers that "we are imagining that these are approved treatments provided by the doctor, and not treatments given during a clinical trial" and asked them to "assume that all your child's medical bills, including the costs of the treatment, are covered by health insurance." We also assured respondents that this "was not a test" and that there were "no right or wrong answers." Each task incorporated a full profile (ie, all 6 attributes were shown) consisting of a specific level for each attribute. Preferences were elicited via caregivers making a judgment as to what aspect constituted the best and then the worst of the treatment.

Survey Instrument

The BWS instrument was included in a broader survey. In addition to basic demographic questions, to ensure that the study sample did not represent individuals with unusually high risk-taking personality traits, the participant section included the 6-item risk-taking measure comprising items from the Jackson Personality Inventory.[46] The previously published mean (SD) score for a physician group was 19 (4), with a range of 11 to 30. We also sought to describe caregivers' numeracy by using the 3-item short form (SNS-3) of the Subjective Numeracy Scale (SNS).[47] Poor numeracy has been associated with ability to evaluate risks and benefits of health options[48] and to negatively affect utility assessment.[49] Previous research has determined a norm mean score for the SNS of 4.03 (1.04), with a range of 1 to 6.[50] Information collected about the affected children included the number of children with DMD, age, where the children lived, a mobility item (with 11 responses ranging from independent walking outdoors to remaining in bed) that represents disease progression, insurance type (private vs government), and whether the children ever experienced a life-threatening emergency.

Survey Pilot: Cognitive Interviewing

Cognitive interviews with 7 parents of individuals with DMD of varying ages and disease stages were used to assess comprehension, refine terminology, and explore the acceptability of the instrument. The investigator (H.L.P.) observed the survey experience of the interviewees by using videoconferencing. Interviewees were asked to "think aloud" as they completed the survey, and the investigator used verbal probes to explore anticipated problem areas, assess understanding of the survey elements, and evaluate willingness to trade among the attributes and levels. The instrument was then modified to the final version based on pilot test feedback, again using a consensus process involving the study team and the advisory group.

Participants and Recruitment

Participants were caregivers (parents or guardians) of at least 1 living child with DMD. The caregivers lived in the United States, were at least 18 years of age, and were able to complete an online survey in English. Their affected child could be any age or at any stage of the disease. The survey was administered online by using the Qualtrics survey system (Qualtrics, LLC, Provo, Utah) from January 17, 2013, to February 21, 2013. In the study advertisements, PPMD committed to sharing the information learned from the survey back to the DMD community. Recruitment occurred with the use of newsletter notices, social media, recruitment e-mails from PPMD, the DuchenneConnect self-report registry, and through word-of-mouth recruiting. The anonymous survey was determined to be exempt by the Western Institutional Review Board (no. 1-756840-1).

Statistical Analysis

In BWS, the dependent variable is the participants' judgment about the best and worst feature in each profile presented to them.[51] Although the results from a BWS can be estimated by using complicated techniques such as conditional logit[52] or hierarchical Bayes,[53] one of the benefits of the method is that it can be analyzed very simply. The simplest techniques focus on the number of times a particular level of an attribute was chosen as best and as worst when it was available in the choice task (unpublished data). A relative best-minus-worst (BW) score can be calculated by subtracting the number of times a feature was chosen as worst from the number of times it was chosen as best, then dividing by the total number of times it was available to be chosen. Early applications of this method have demonstrated a very high level of correlation between such simple techniques and more complicated regression-based techniques.[39,40] As with all techniques used to estimate ordinal, multinomial

outcomes, our scoring approach assumes equal spacing between things that were chosen as best (BW score, 1), those that were not chosen (BW score, 0), and those chosen as worst (BW score, –1).

We chose to estimate the importance weights for each level by using the relative BW score because it could be easily understood by the broadest readership (including the community of patients and caregivers that we engaged in developing this survey). In addition to the simplicity of the BW score, this approach has several advantages. First, regression-based techniques require the use of the omitted category[39] or the use of complex effects coding procedures[40] to estimate choice models. Second, by using this simple approach, we have ensured that all estimated parameters remain on the same ratio scale. This method allows comparisons to be made across the attributes, as well as identification of global best and worst attribute levels across all the attributes. Because the BW score is estimated as a mean across the sample, we also report the SEs for these means. This process allows us to conduct *t* tests to determine whether the scores were significantly different from zero. We did this for each attribute level and have reported the *P* values for each test.

Finally, we used the relative BW score for each level within each attribute to assess the overall importance of each attribute, conditioned on the levels chosen.[54] With this technique, relative attribute importance was estimated by subtracting the lowest relative BW score associated with a level of that attribute from the highest relative BW score associated with a level of that attribute. We then divided each difference by the sum of all differences across the 6 attributes and reported the result as a percentage.

RESULTS

A total of 124 caregivers who self-identified as being a parent or guardian of an individual with DMD began the treatment experiment. Two individuals dropped out after the first treatment task; 1 dropped out after the third treatment task; 1 dropped out after the fifth treatment task; and 1 dropped out after the 15th treatment task. The remaining 119 caregivers completed the entire survey.

Table II summarizes the characteristics of the sample. Caregivers were predominantly white, married, biological mothers, and had 1 affected child. Education level ranged from high-school or General Educational Development diploma to graduate or professional degree; the median response was "4-year college degree." Annual household income ranged from "<$25,000" to ">$100,000"; the median response was "$75,000 to $100,000." Caregivers' ages ranged from 28 to 66 years (mean, 43.7 years), and the age of the affected children ranged from 2 to 38 years (mean, 12.1 years). Slightly more than one half of the children were reported as having participated in clinical research and more than one third in a clinical trial. Almost all of the affected children lived in the home of the caregiver (n = 117 [98%]). The majority of the affected children used private insurance for their medical care (n = 101 [85%]), although 34% (n = 40) endorsed that their child used a state/government program. Caregivers reported that 19% (n = 22) of their children had experienced a life-threatening emergency.

The affected children represented a range of disease progression. When these children were dichotomized into an "ambulatory" group, defined as those who could walk independently outdoors for at least short distances, and a "nonambulatory" group, defined as those who could not walk outdoors without help, 75 (64%) were parents with children in the ambulatory group and 43 (36%) were parents with children in the nonambulatory group.

The mean (SD) risk-taking score was 17 (4), with a range of 7 to 30 (higher scores indicate more risk-taking endorsement). The caregivers in this study scored significantly lower on the risk-taking score than the physician reference group[41] (mean, 19 [4]; $P < 0.005$), indicating lower risk-taking personality traits. The mean SNS-3 score was 4.90 (1.1), with a range of 2.33 to 4.87, which was higher than the reference population.[42] This result was consistent with the high educational levels reported in our study (68% with at least a college degree).

Table III presents relative BW scores, SEs, and *P* values, and Figure 2 diagrams the relative utility of each level as measured by using relative BW scores. All of the BW scores were significant at $P < 0.001$ except for "no increased chance of nausea," "no postapproval drug information available," and "no increased risk of bleeding." By a large margin, the highest utilities as measured by using relative BW scores were for "stops progression of weakness" (0.877) and "slows progression of weakness" (0.800). These scores had almost twice the utility of

Table II. Characteristics of participants and affected children (N=119)

Characteristic	Value
Participant characteristics	
Mean (SD) caregiver age, y	43.7 (7.7)
Mean (SD) child age, y	21.1 (6.4)
Caregiver characteristics	
Relationship to children	
Biological mother	67.2%
Biological father	28.6%
Adoptive mother	3.4%
Adoptive father	0.8%
Marital status	
Married/long-term relationship	89.9%
Divorced/separated	9.2%
Widowed	0.8%
Race	
White	91.6%
Education	
High school/GED	4.2%
Some college	14.3%
Technical school	5.0%
Associated degree	7.6%
4-year college degree	42.9%
Graduate/professional degree	25.2%
Income	
<$25,000	5.9%
$25,000–$50,000	8.4%
$50,000–$75,000	18.5%
$75,000–$ 100,000	18.5%
>$100,000	47.1%
Child characteristics	
No. of affected children	
1 child	92.4%
≥2 children	7.6%
Living arrangements	
In caregiver's home	98.3%
Independent	0.8%
Other	0.8%
Ambulation status	
Ambulatory*	63.9%
Nonambulatory	36.0%
Research participation	
Clinical research	58.0%
Clinical trial	34.0%
Had life-threatening emergency	
Yes	18.5%
No	81.5%

Note: In some cases, percents do not add to 100% because of missing values.
GED = General Educational Development diploma.
*Ability to walk independently outside for at least short distances.

the next-highest score, "5-year gain in expected life span" (0.464). The "2-year gain in expected life span" had a similar priority to the "5-year gain in expected life span" (0.408).

Caregivers attributed the highest negative BW scores to "increased risk of dangerous heart arrhythmia and sudden death" (–0.786), followed by "increased risk of hemorrhagic stroke and lifelong disability" (–0.720). This was followed by "causes loss of appetite with occasional vomiting" (–0.280). Although the 2 most serious risks had high negative scores, either (but not both) could be offset by a treatment that stopped the progression of weakness. The amount of knowledge about the drug was not given high relative BW scores at any level, with mean scores ranging from 0.056 to –0.021.

Table IV includes the relative attribute importance for the entire group of caregivers. At the attribute level, effects on muscle function accounted for the largest proportion of the variance (28.7%), followed by arrhythmia (22.4%), bleeding (21.2%), life span (17.3%), nausea (8.1%), and knowledge about the drug (2.3%).

DISCUSSION

Although the FDA is committed to patient-centered drug development, the agency has limited resources. Representing a disease community that was not selected for the congressionally mandated community engagement program, PPMD led a study to proactively inform the FDA's benefit–risk assessments. The process we used can be a model for facilitating patient-centered drug development through an exploration of the priorities and

Table III. Best-worst results.

Attribute description	Best	Worst	Best-worst	Relative b-w Score	S.E.	T-test	P-Value
Effect on muscle function							
Stops the progression of weakness	628	2	626	877	0.013	69.441	<0.001
Slows the progression of weakness	571	0	571	0.800	0.015	53.357	<0.001
Does not change progression of weakness	68	125	−57	−0.080	0.019	−4.149	<0.001
Lifespan							
5 veargain in expected lifespan	348	17	331	0.464	0.020	22.741	<0.001
2 veargain in expected lifespan	299	3	291	0.408	0.019	21.186	<0.001
No extra gain to expected lifespan	12	93	−81	−0.113	0.014	−8.269	<0.001
Knowledge about the drug							
2 years of post-approval drug info available	109	69	40	0.056	0.019	3.015	0.001
1 year of post-approval drug info available	20	4	16	0.022	0.007	3.288	0.001
No post-approval drug info available	41	56	−IS	−0.021	0.014	−1.524	0.064
Nausea							
No increased chance of nausea	19	26	−7	−0.010	0.009	−1.044	0.148
Causes loss of appetite	1	95	−94	−0.132	0.013	−10.272	<0.001
Causes loss of appetite with occasional vomiting	17	217	−200	−0.280	0.019	−14.981	< 0.001
Risk of bleeds							
No increased risk of bleeds	3	11	-8	−0.011	0.005	−2.143	0.016
Increased risk of bleedinggums and increased bruising	0	190	−190	−0.266	0.017	−16.079	<0.001
Increased risk of hemorrhagic stroke and lifelong disability	0	514	−514	−0.720	0.017	−42.807	< 0.001
Risk of heart arrhythmia							
No increased risk of heart arrhythmia	5	32	−27	−0.038	0.008	−4.498	<0.001
Increased risk of harmless heart arrhythmia	1	122	−121	−0.169	0.014	−11.943	<0.001
Increased risk of dangerous arrhythmia and sudden death	0	551	−561	−0.786	0.015	−51.131	<0.001

preferences of patients, families, and other stakeholders. Although the individual stories of highly motivated advocates are powerful and influential, it is difficult to know whether these testimonials represent the perspectives of the majority of patients and families. We describe a successful community-engaged process to understand treatment preferences in a large group of decision makers supported by the use of best/worst scaling. To the best of our knowledge, this study represents the first time a patient advocacy organization has led a quantitative preferences study of this complexity, highlighting a successful advocacy/academic collaboration that integrates traditional advocacy methods, family-centered outcomes research, and a scientific approach to quantifying preferences.

Within the context of our experiment, caregivers attributed very high scores to stopping or slowing the progression of muscle weakness. Change in life span was not scored as highly. Feedback during cognitive interviewing suggested that parents associated better muscle function with higher quality of life, indicating that parents value quality more than length of life. This finding is consistent with both anecdotal reports and an interview study of parents of children involved in clinical trials,[41] in which parents expressed a preference for better quality of life for their child over a longer life span.

We found that the presence of a serious risk could be compensated for by a treatment that stops or slows progression of weakness, even absent any other benefits. The burden of DMD may be associated with parents' willingness to accept a serious risk for a noncurative treatment. The data support a limit to parents' risk tolerance, however: for the levels of benefit provided in the experiment, they would not accept a treatment with 2 serious risks.

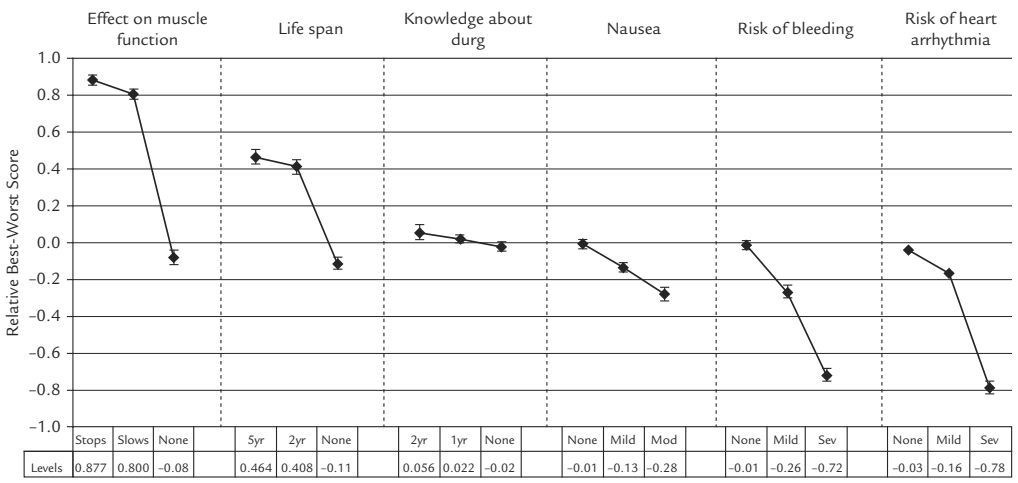

Figure 2. Relative best–worst scores for attribute levels. Mod = moderate; Sev = severe.

Our community-engaged process contributed to successful recruitment of sufficient numbers of caregivers for a complex, time-intensive survey in only 5 weeks, notwithstanding the fact that the study focuses on a rare disease. The caregivers' children represented a range of ages and disease stages, and thus our outcomes reflect the preferences of parents with children across the disease course. Although the development of an appropriate experimental design is a complex task, it is one that is well suited to be led by advocacy organizations with expert input and collaboration.

Limitations

There are several limitations to the study. First, the study sample, although likely to be representative of caregivers whose children are enrolled in clinical trials, may not be generalizable to the broader DMD community. However, we have demonstrated that this population was not unusually high in risk-taking

Table IV. Relative attribute importance.

Attribute	Maximum	Minimum	Difference	Percent*
Effect on muscle function	0.877	−0.080	0.957	28.7
Life span	0.464	−0.113	0.577	17.3
Knowledge about the drug	0.056	−0.021	0.077	2.3
Nausea	−0.010	−0.280	0.270	8.1
Risk of bleeding	−0.011	−0.720	0.709	21.2
Risk of heart arrhythmia	−0.038	−0.786	0.748	22.4
Sum			3.338	100

*Percent relative importance calculated as the difference between the maximum and minimum utility for each attribute divided by the sum of all such differences.

personality traits and had adequate numeracy to reduce concern about numeracy bias in survey responses. Although the recruitment of caregivers (or patients) through advocacy groups has a risk of bias, it also has real benefits over qualitative approaches. Using the model process, we plan to refine the experiments and conduct a larger study with a more representative parent group and a neuromuscular clinician group. Especially important in our next study is to elicit treatment preferences from affected teenagers and adults, anticipating that DMD patients and caregivers may not assess benefit and risk in the same way.[19]

Second, although the study used a rigorous approach to attribute identification, the simulated treatments described in the experiment may not represent the benefit and risk profile of therapies that are ultimately approved for DMD. As with all stated-preference experiments, it remains unknown whether the presence or absence of additional attributes would influence the results. On the spectrum of patient centered to clinically centered specification of attributes, we favored the former to be consistent with the goals of patient-centered outcomes research and explore attributes meaningful to our caregiver participants. In our future studies, and when more is known about the benefits and risks of treatment, we aim to incorporate more clinically centered attributes while continuing to maintain a priority on utilizing meaningful attributes.

Third, we conducted an aggregate analysis, and important structures in preference heterogeneity may have been overlooked. We have previously reported the differences in treatment preferences by stratifying data according to child's ambulation status.[55] Although there was a small but significant difference when completing such a stratification, this could have been explained by scale differences between the 2 groups. In follow-up studies, we will aim to have a larger sample size to allow for both stratification and segmentation analysis,[56] which will enable being able to adequately describe preference heterogeneity.

Fourth, we used a simple technique for estimating preferences, compared with more advanced regression techniques. As a supplemental analysis (not reported here), we reanalyzed our data by using a conditional logit. One obvious difference between the methods is that conditional logit requires using effects coding for each attribute, making each attribute have the same mean. As such, although each attribute remains on a ratio scale, the translocation of the origin inherent in effects coding implies that level importance cannot be compared across attributes. The advantage of our simple approach is that all preference weights can be estimated directly (ie, without using effects coding), and hence they all sit on the same ratio scale. We modified these results to make them comparable to the conditional logit (ie, we subtracted the attribute mean from each attribute level), and they produced nearly identical results to the conditional logit, with both methods having identical ordering (Spearman's $\rho = 1.0$) and near perfect correlation (Pearson's $\rho = 0.997$).

Finally, because BWS is a relatively new stated-preference method, there is the possibility that it may present a distorted version of preference. However, there is growing interest in the method given its simplicity compared with more traditional conjoint analysis methods, which may affect respondent efficiency (ie, do responses to choice tasks reflect respondents true preferences?). We plan to validate these results against a simple conjoint analysis that was conducted as part of this study, but more research is needed to compare BWS and conjoint analysis methods.

Implications

The study findings are highly relevant to industry and regulators who are conducting benefit–risk assessments for potential DMD therapies. Emerging results from clinical trials suggest a slowing of motor decline, as measured by using the 6-minute talk test, and no known effect on life span.[13] Caregivers' significant and yet finite risk tolerance has regulatory implications as well; however, given the modest risk profile emerging from many DMD clinical trials, our finding of high tolerance for adverse effects and drug-related uncertainty is also relevant.

This study intended to leverage the FDA's ongoing commitment to identifying methods of systematic patient engagement and, more specifically, their commitment to the use of statistical methods exploring and comparing benefits and risks to systematically quantify patients' anecdotal reports.[57] PPMD was able to report the outcome of this study to FDA representatives in both private and public meetings. Equally important, PPMD reported the results back to the DMD community through social media, a webinar, and in-person meetings and conferences. As the FDA evaluates new drug applications for DMD therapies, they should be mindful of the value that

parent decision makers place on even moderate benefits to function, their tolerance for considerable risk, and their tolerance for uncertainty.

ACKNOWLEDGMENTS

We appreciate the leadership and commitment of the PPMD oversight committee: Pat Furlong, Brian Denger, Sharon Hesterlee, and Kathi Kinnett. We are indebted to the stakeholder informants, parents who participated in the cognitive interviews, and caregivers who completed the survey.

Ms. Peay was responsible for community engagement, survey development, and writing and contributed to data analysis and data interpretation. Ms. Hollin was responsible for figure creation and contributed to data analysis, data interpretation, and writing. Mr. Fischer contributed to community engagement and survey development. Dr. Bridges was responsible for study design, data analysis, and data interpretation and contributed to survey development and writing.

CONFLICTS OF INTEREST

This study was funded by Parent Project Muscular Dystrophy, which had a role in the study design; in the collection, analysis, and interpretation of data; in writing the manuscript; and in the decision to submit the manuscript for publication. Ms. Peay and Mr. Fischer are employees of PPMD, and Dr. Bridges received compensation from PPMD for this project. The authors indicated that they have no conflicts of interest regarding the content of this article.

REFERENCES

1. Centers for Disease Control and Prevention (CDC). Prevalence of Duchenne/Becker muscular dystrophy among males aged 5-24 years—four states, 2007. *MMWR Morb Wkly Rep*. 2009;58:1119–1122.
2. Emery AE. Population frequencies of inherited neuromuscular diseases—a world survey. *Neuromuscul Disord*. 1991;1:19–29.
3. Bushby K, Finkel R, Birnkrant DJ, et al. Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and pharmacological and psychosocial management. *Lancet Neurol*. 2010;9:77–93.
4. Eagle M, Baudouin SV, Chandler C, et al. Survival in Duchenne muscular dystrophy: improvements in life expectancy since 1967 and the impact of home nocturnal ventilation. *Neuromuscul Disord*. 2002;12:926–929.
5. Pangalila RF, van den Bos GA, Stam HJ, et al. Subjective caregiver burden of parents of adults with Duchenne muscular dystrophy. *Disabil Rehabil*. 2012;34:988–996.
6. Kenneson A, Bobo JK. The effect of caregiving on women in families with Duchenne/Becker muscular dystrophy. *Health Soc Care Community*. 2010;18:520–528.
7. Larkindale J, Yang W, Hogan PF, et al. Cost of illness for neuromuscular diseases in the US. *Muscle Nerve*. 2014; 49:431–438.
8. Ouyang L, Grosse SD, Fox MH, Bolen J. A national profile of health care and family impacts of children with muscular dystrophy and special health care needs in the United States. *J Child Neurol*. 2012;27:569–576.
9. Ouyang L, Grosse SD, Kenneson A. Health care utilization and expenditures for children and young adults with muscular dystrophy in a privately insured population. *J Child Neurol*. 2008;23:883–888.
10. Biggar WD, Harris VA, Eliasoph L, Alman B. Long-term benefits of Deflazacort treatment for boys with Duchenne muscular dystrophy in their second decade. *Neuromuscul Disord*. 2006;16:249–255.
11. Angelini C. The role of corticosteroids in muscular dystrophy: a critical appraisal. *Muscle Nerve*. 2007;36:424–435.
12. Bushby K, Muntoni F, Urtizberea A, et al. Report on the 124th eNMC International workshop. Treatment of Duchenne muscular dystrophy; defining the gold standards of management in the use of corticosteroids. *Neuro Muscul Disord*. 2004;14:526–534.
13. Govoni A, Magri F, Brajkovic S, et al. Ongoing therapeutic trials and outcome measures for Duchenne muscular dystrophy. *Cell Mol Life Sci*. 2013;70:4585–4602.
14. Milne CP. Prospects for rapid advances in the development of new medicines for special medical needs. *Clin Pharmacol Ther*. 2014;95:98–109.
15. Michaels M, Collyar D. Research advocacy in traditional research settings: questions of influence and legitimacy. In: Earp JL, French EA, Gilkey MB, eds. *Patient Advocacy for Healthcare Quality: Strategies for Achieving Patient-Centered Care*. Boston, Mass: Jones and Bartlett; 2007:445–473.
16. Parent Project Muscular Dystrophy. Benefit-risk assessments in rare disorders: The case for therapeutic development in Duchenne muscular dystrophy as the prototype for new approaches. http://www.parentprojectmd.org/BR. Accessed April 28, 2014.
17. Genetic Alliance UK. New medicines for serious conditions: weighing the risks and benefits. The verdict of a jury of patients. http://www.geneticalliance.org.uk/docs/citizens-jury-report.pdf. Accessed March 1, 2014.
18. Parent Project Muscular Dystrophy: Share Your Brave Story. http://community.parentprojectmd.org/profiles/blog/list?tag=My+BRAVE+Story. Accessed March 1, 2014.
19. McNeil DE, Davis C, Jillapalli D, et al. Duchenne muscular dystrophy: drug development and regulatory considerations. *Muscle Nerve*. 2010;41:740–745.
20. Hauber B, Fairchild AO, Johnson FR. Quantifying benefit-risk preferences for medical interventions: an overview of

a growing empirical literature. *Appl Health Econ Health Policy*. 2013;11: 319–329.

21. US Food and Drug Administration (FDA). Enhancing benefit-risk assessment in regulatory decision-making. http://www.fda.gov/forindustry/userfees/prescriptiondrugu serfee/ucm326192.htm. Accessed March 1, 2014.
22. Mullard A. Patient-focused drug development programme takes first steps. *Nat Rev Drug Discov*. 2013;12: 651–652.
23. Facey K, Boivin A, Gracia J, et al. Patients' perspectives in health technology assessment: a route to robust evidence and fair deliberation. *Int J Technol Assess Health Care*. 2010;26:334–340.
24. Hansen HP, Draborg E, Kristensen FB. Exploring qualitative research synthesis: the role of patients' perspectives in health policy design and decision making. *Patient*. 2011; 4:143–152.
25. Vogt F, Schwappach D, Bridges J. Accounting for tastes: a German perspective on the inclusion of patient preferences in health care. *Pharmacoeconomics*. 2006;24:419–423.
26. Coplan P, Noel R, Levitan B, et al. Development of a framework for enhancing the transparency, reproducibility and communication of the benefit–risk balance of medicines. *Clin Pharmacol Ther*. 2010; 89:312–315.
27. US Food and Drug Administration (FDA). *Factors to Consider When Making Benefit-Risk Determinations in Medical Device Premarket Approval and De Novo Classifications. FDA Guidance for Industry and Food and Drug Administration Staff*. Silver Spring, Md: Center for Devices and Radiological Health, Center for Biologics Evaluation and Research. Food and Drug Administration. March 28, 2012.
28. Fleurence R, Selby JV, Odom-Walker K, et al. How the Patient-Centered Outcomes Research Institute is engaging patients and others in shaping its research agenda. *Health Affairs*. 2013;32:393–400.
29. Bridges J, Hauber AB, Marshall D, et al. A checklist for conjoint analysis applications in health: report of the ISPOR Conjoint Analysis Good Research Practices Taskforce. *Value Health*. 2011;14:403–413.
30. Marshall D, Bridges J, Hauber AB, et al. Conjoint analysis applications in health—how are studies being designed and reported? An update on current practice in the published literature between 2005 and 2008. *Patient*. 2010;3:249–256.
31. Bridges J, Kinter E, Schmeding A, et al. Can patients diagnosed with schizophrenia complete choice-based conjoint analysis tasks. *Patient*. 2011; 4:267–275.
32. Coast J, Flynn TN, Salisbury C, et al. Maximising responses to discrete choice experiments: a randomised trial. *Appl Health Econ Health Policy*. 2006;5:249–260.
33. Reed Johnson F, Lancsar E, Marshall D, et al. Constructing experimental designs for discrete-choice experiments: report of the ISPOR Conjoint Analysis Experimental Design Good Research Practices Task Force. *Value Health*. 2013;16: 3–13.
34. International Society for Pharmacoeconomics and Outcomes Research. ISPOR Statistical Methods for Conjoint Analysis Task Force. http://www.ispor.org/TaskForces/Analyzing_Data_Conjoint_Analysis.asp. Accessed March 1, 2014.
35. Erdem S, Rigby D, Wossink A. Using best-worst scaling to explore perceptions of relative responsibility for ensuring food safety. *Food Policy*. 2012;37:661–670.
36. Flynn TN, Louvier JJ, Peters TJ, Coast J. Best-worst scaling: what it can do for health care research and how to do it. *J Health Econ*. 2007; 26:171–189.
37. Bridges J. Stated preference methods in health care evaluation: an emerging methodological paradigm in health economics. *Appl Health Econ Health Policy*. 2003;2:213–224.
38. Flynn TN. Valuing citizen and patient preferences in health: recent developments in three types of best-worst scaling. *Expert Rev Pharmacoecon Outcomes Res*. 2010;10:259–267.
39. Louviere JJ, Flynn TN. Using best-worst scaling choice experiments to measure public perceptions and preferences for healthcare reform in Australia. *Patient*. 2010;3:275–283.
40. Gallego G, Bridges J, Flynn T, et al. Using best-worst scaling in horizon scanning for hepatocellular carcinoma technologies. *Int J Technol Assess Health Care*. 2012;28:339–346.
41. Peay HL, Sheffer H, Tibben A. Expectations and decision making in clinical trials for Duchenne and Becker muscular dystrophy. Abstract presented at: 18th International Congress of The World Muscle Society; October 1–5, 2013; Asilomar, Calif.
42. National Center for Education Statistics. National assessment of adult literacy (NAAL). A first look at the literacy of America's adults in the 21st century. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2006470. Accessed March 1, 2014.
43. Townsend D, Yasuda S, Li S, et al. Emergent dilated cardiomyopathy caused by targeted repair of dystrophic skeletal muscle. *Mol Ther*. 2008; 16:832–835.
44. Shirokova N, Niggli E. Cardiac phenotype of Duchenne muscular dystrophy: insights from cellular studies. *J Mol Cell Cardiol*. 2013;58: 217–224.
45. Kuhfeld W. Orthogonal arrays [TS-723]. Cary (NC): SAS http://support.sas.com/techsup/technote/ts723.html. Accessed March 1, 2014.
46. Pearson SD, Goldman L, Orav EJ, et al. Triage decisions for emergency department patients with chest pain: do physicians' risk attitudes make the difference? *J Gen Intern Med*. 1995;10:557–564.

47. Wallston KA, McNaughton C, Storrow A, et al. Validation of a short, 3-item version of the Subjective Numeracy Scale (SNS-3). Presented at: Health Literacy Research Conference; October 17-18, 2011; Chicago, IL.
48. Peters E, Hibbard J, Slovic P, Dieckmann N. Numeracy skill and the communication, comprehension, and use of risk-benefit information. *Health Affairs*. 2007;3:741–748.
49. Zikmund-Fisher B, Smith DM, Ubel PA, Fagerlin A. Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. *Med Decis Making*. 2007;27:663–671.
50. Fagerlin A, Zikmund-Fisher BJ, Ubel PA, et al. Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Med Decis Making*. 2007;27:672–680.
51. Molassiotis A, Emsley R, Ashcroft D, et al. Applying best-worst scaling methodology to establish delivery preferences of a symptom supportive care intervention in patients with lung cancer. *Lung Cancer*. 2012;77:199–204.
52. McFadden D. Conditional logit analysis of qualitative choice behavior. In: Zarembka P, ed. *Frontiers in Econometrics*. New York, NY: Academic Press; 1974:105–142.
53. Orme BK. *Getting Started With Conjoint Analysis: Strategies for Product Design and Pricing Research*. *Research Publishers*. 2006.
54. Bridges J, Mohamed AF, Finnern HW, et al. Patients' preferences for treatment outcomes for advanced non-small cell lung cancer: a conjoint analysis. *Lung Cancer*. 2012;77:224–231.
55. Peay HL, Hollin I, Sheffer H, Bridges J. Measuring treatment preferences of children with Duchenne muscular dystrophy using best-worst scaling. Presented at the 63rd Annual Meeting of the American Society of Human Genetics; October 23, 2013; Boston, Mass.
56. Deal K. Segmenting patients and physicians using preferences from discrete choice experiments. *Patient*. 2014;7:5–21.
57. Johnson FR, Hauber AB, Poulos CM. Research Report: A Brief Introduction to the Use of Stated-Choice Methods to Measure Preferences for Treatment Benefits and Risks. RTI Press publication No. RR-0009-0909. Research Triangle Park, NC: RTI International. http://www.rti.org/rtipress. Accessed December 20, 2013.

Address correspondence to: Holly L. Peay, MS, Parent Project Muscular Dystrophy, 401 Hackensack Avenue, 9th Floor, Hackensack, NJ 07601. E-mail: holly@parentprojectmd.org