Clinical outcome measures for trials in Duchenne muscular dystrophy: report from International Working Group meetings

Clin. Invest. (2011) 1(9), 1217–1235

In June 2010, 25 representatives from Europe and the US met in Washington, DC, USA, to discuss clinical outcome measures in Duchenne muscular dystrophy (DMD) in the context of clinical trial design and analysis. The workshop was organized in response to a September 2009 European Medicines Agency meeting where a clear directive was given that an international consensus needs to be developed that provides a foundation for age-appropriate clinical outcome measures for use in clinical trials of emerging therapeutics for DMD. Data were presented from eight multicenter longitudinal datasets, representing nearly 1900 patients over a 20-year time period. This experience confirmed the feasibility of repeated evaluations performed at multiple sites and addressed several core issues in drug development for DMD, such as the 'new' natural history in the steroid-era, reliability and sensitivity of specific outcome measures, as well as disease staging and patient selection. These data form a valuable asset for academic investigators, pharmaceutical sponsors and regulatory agencies involved in DMD therapeutics. The group remains committed working together on a number of collaborative goals to support the therapeutics development effort in this orphan disease and to make these data available to stakeholders working in the field.

 Keywords: 6-min walk test • clinical outcome measures • clinical trial design • Cooperative International Neuromuscular Research Group
 Duchenne muscular dystrophy • Medical Research Council score • natural history of Duchenne muscular dystrophy • North Star Ambulatory Assessment
 • pulmonary function tests • quantitative muscle testing • TREAT-NMD

The International Duchenne muscular dystrophy Clinical Outcomes Working Group

During the past decade significant progress has been made in development of candidate drugs for the treatment of Duchenne muscular dystrophy (DMD). Perhaps the most notable of these are antisense oligonucleotides for exon skipping [1]. As candidate drugs enter clinical trials, it is critical to understand the contemporary natural history of DMD and utilize available data from natural history studies to evaluate clinical outcome measures for planned efficacy studies. On 28–29 June 2010, 25 participants from Europe and the US met in Washington, DC, USA, to discuss clinical outcome measures in DMD in the context of clinical trial design. The workshop was organized by TREAT-NMD and Children's National Medical Center and was supported by Cure Duchenne, the Foundation to Eradicate Duchenne and Ryan's Quest. A followup meeting was held in Naples on 21 July 2010. Participants included representatives of international studies of DMD natural history and clinical trials experts.

Kate Bushby^{†1} & Edward Connor²

¹Institute of Genetic Medicine, International Centre for Life, Newcastle University, Newcastle upon Tyne, UK ²Department of Pediatrics, Clinical & Translational Research Institute, Children's National Medical Center, Washington, DC, USA 'Author for correspondence: E-mail: kate.bushby@ncl.ac.uk *For the International DMD Clinical Outcomes Working Group. A list of all members of the working group can be found at the end of this article.



The clinical outcomes workshops were held in response to a September 2009 European Medicines Agency meeting [2], where a clear directive was given that an international consensus needs to be developed that provides guidance on age-appropriate clinical outcome measures for use in clinical trials for DMD. Although measures of motor function, such as timed tests, functional scales and the 6-min walk test (6MWT) have been validated and used in recent or ongoing clinical trials, it was felt that data were needed on the relationships of these measures to the clinical progression of DMD in the context of the contemporary natural history of the disease. In addition, there was widespread recognition that the commonly used measures did not address the varying levels of functional ability that are seen in patients with DMD at different ages. A strong understanding of the natural history of the disease is essential to drug development in orphan diseases similar to DMD, since the numbers of patients available for trials will be limited.

The aims of the meeting were to:

- Map the outcome measures currently used in DMD natural history studies and clinical trials, and to highlight and assess the data currently being collected;
- Establish the current expected clinical course of the disease from these contemporary datasets;
- Determine whether data exist to define the relationship between the outcome measures and milestones of disease progression;
- Determine areas where the available data clearly describe the disease and where gaps in data highlight the need for further work.

Preliminary work

To establish the international dataset on which to base the discussion, preliminary work identified eight natural history and clinical trial datasets, representing over 1900 patients who are being, or have been, prospectively followed. This work took place over the winter and spring of 2009–2010.

The demographics of participants in these datasets are shown in Table 1. The various measures included in each dataset were compiled to see where the most data were concentrated and identify gaps, such as the age groups or the range of investigations covered (Supplementary Table 1).

Clinical course & progression of the disease

Investigators from each participating group were asked to present their data in a format that facilitated mapping disease progression and to focus on those measures that were most commonly collected. A guideline was developed to aid in data analysis and presentation (Box 1). Outcome measures were described in terms of their sensitivity, reliability and applicability in clinical trials as well as relationship to disease progression. A facilitated discussion followed the presentations. A summary of key data presented from each dataset follows.

Meeting proceedings: presentation of the available data

Longitudinal assessment in confirmed DMD: Cooperative International Neuromuscular Research Group Study

CM McDonald, RT Abresch, EK Henricson

A detailed overview of specific outcomes from this multicenter international study conducted by the Cooperative International Neuromuscular Research Group (CINRG) group was provided. The presentation included discussion of details regarding study methodology, optimal methods for evaluating timed motor performance data, 1-year follow-up of ranges and variability in strength, timed motor performance and pulmonary function measures, and functional scale and timed motor performance data as predictors of loss of functional 'milestones'.

Study methodology

The study aims included:

- Producing a high-quality 'steroid-era' contemporary dataset for use by clinical researchers;
- Understanding the interrelation of quantitative measures of disease and qualitative reports of subjective well-being.

The study includes 340 individuals with confirmed DMD between the ages of 2 and 28 years being followed at 20 participating CINRG centers in eight countries. Participants are assessed at months 3, 6, 9, 12 and annually thereafter to 5 years. Assessments include evaluation of strength (Medical Research Council [MRC] manual muscle test, quantitative strength testing) [3,4] timed motor performance testing (time to stand from supine, time to climb four standard stairs, time to walk/run 10 m)[3], Brooke and Vignos limb function scales [3,5] and pulmonary function testing (forced vital capacity [FVC], forced expiratory volume in 1 s, peak expiratory flow rate, maximal inspiratory and expiratory pressures), and health-related quality of life. Data presented included follow-up through the end of the first year of study participation, and participants were stratified by year of age, glucocorticoid treatment status (current vs previously treated or naive), and ability to walk independently as judged by the site investigator.

Clinical outcome measures for trials in Duchenne muscular dystrophy Meeting Proceedings

Table 1. Dataset demographic	:S.			
Database	Patients (n)	Number steroid positive (age spread)	Number steroid negative (age spread)	Ambulant/non-ambulant
Italian DMD Telethon Network	106	96 (4.1–17)	10 (4–9.3)	Ambulatory
North Star UK	300 (240 included in analysis)	223 (3–16)	17 (3–13)	Mainly ambulatory- non-ambulant cohort beginning to be collected
Cyclosporine A study, MD-NET	146	146 (4.5–11.7)		Ambulatory
Ataluren (PTC124) study 007	57 (placebo arm)	40 (5.3–15.3)	17 (5.1–14.0)	Ambulatory
CINRG	348	Current treatment at baseline 210 (2–28) Prior treatment >1 month at baseline 48 (8–28)	82 (2–27)	Both
MFM	1000 patients (163 are DMD)	27 (6–27)	136 (4–33)	Both
Denmark	160 patients 100 >15 years 60 <15 years	50 (5–15)	110 (5–48)	Both
United Dystrophinopathy Project	776 DMD patients (671 in analysis) + 177 BMD 46 IMD	500 (2.9–34.3)	171 (0.5–37.8)	Both
Total	1996			
BMD: Becker muscular dystrophy; DMD: Duchenne muscular dystrophy; IMD: Intermediate muscular dystrophy.				

Interpretation of timed motor performance data Timed motor performance evaluations (time to rise from supine, time to climb four standard stairs, and time to walk or run 10 m) are commonly used as clinical trial outcome measures, are reliable measures of functional ability and have been used widely in DMD clinical trials [3,4,6]. However, efforts to evaluate data in populations where some participants have lost ambulation have

Box 1. Guidance.

- Basic demographics of patient population and study design.
- How does your data define progression of the disease?
- Do the outcome measures (OMs) you collected reflect the natural history of the disease?
- Could any be used as a surrogate or proxy measure for a milestone/event?
- How well does it track progression of the disease (i.e., sensitivity)?
- Could your data be applied to predict a milestone or event? For example would it be possible to say that scoring 'x' at a particular time point means that 'y' will occur within a specific time frame?
- Do any of the OMs you collected correlate with any other collected OMs?
- Are there other disease milestones that could be documented more sensitively or are better at predicting disease progression? Can this inform biomarker projects?
- What is the variability between observers of your outcome measures? That is, inter- and intra-rater variability?
- Are they reliable for use in multicenter clinical trials? Or are they difficult to reproduce outside of a single center?
- What is the level of heterogeneity of measures across different age groups? Is a measure more or less variable in different age groups or at different functional stages?
- Is there information on variability across the population?
- Are there clinically important areas we are not capturing (e.g., energy levels)?

been hampered by the fact that the inability to complete an evaluation leads to a theoretically infinite maximum time, which cannot be used in statistical analyses. The alternative is to set an arbitrary maximum allowable time score such as 120 s to complete a 10-m walk.

Data from a group of DMD boys and similar aged controls is shown in Figure 1, but this method creates a false appearance of a linear relationship in the data that is difficult (if not impossible) to interpret because the ceiling score does not represent an actual value of infinite time. Because of this challenge, interpretation of timed motor performance testing has been limited to somewhat homogeneous groups who are able to complete the testing, thus creating results that are biased in favor of the more functional individuals. However, this can be avoided by converting time scores into velocities. For example (Figure 2), when the same 10-m walk data is converted to a walking velocity (in m/s) this results in a linear pattern of decline that adequately represents the impact of the 'zero velocities' of individuals who are unable to perform the evaluation. This representation of the data is simple to interpret and facilitates straightforward comparisons between the DMD and control groups.

Reliability, feasibility & variability of outcome measures

Measures of strength, timed motor performance and pulmonary function have been well described and validated in DMD patient populations [4,6,7], but there are design limitations that restrict their feasibility and/or



Figure 1. 10-m run/walk times in Duchenne muscular dystrophy patients and controls by age with an arbitrary 120 s maximum for individuals who are unable to perform the task. The slope of the line for DMD participants is not representative of a true rate of change and is influenced by the artificial scores for non-ambulatory participants.

DMD: Duchenne muscular dystrophy.

utility at some ages and stages of disease in ways that might place design limitations on future clinical trials. Key considerations include the ages at which measures can detect differences between steroid-treated and naive patients, where they reflect representative age groups versus subpopulations thereof, ability to detect agerelated changes, and variability and reliability of measures over time. All of these factors impact a given tool's utility as primary clinical trial end points.

Several conclusions are possible based on analysis of data from the current CINRG longitudinal study cohort. First, strength testing is reliable and reflects differences between steroid-treated and naive populations between the ages of 4 and 9 years and for stronger and more mobile subpopulations aged 10 and older. However, strength testing has limited continuity across the entire age range of affected individuals from young children to adults. Currently available methods show significant 1-year decreases in mean muscle testing scores for children in the 7-9 and 10-12 year-old age group, but not for those 13-18 years of age. Quantitative strength measures revealed significant growth-related increases in strength in the knee and elbow flexors over a 1-year time period in the 4-6 year age group. However, progression of the disease is evidenced by the significant 1-year loss of strength for the knee flexors and elbow flexors and extensors in the 7-9, 10-12 and 13-18 year age groups. Knee extensors revealed significant 1-year decline in the 7–9 and 10–12 year-old age groups, but not 13–18 year olds. Grip strength significantly increased over a 1-year period in the 4-6, 7-9, and 10-12 year age groups, exhibited no change in the 13-18 year age group, and showed a significant loss in the >18 year-old group. However, this needs to be interpreted with caution, as the data also include the normal growth-associated increases in strength. Data were presented suggesting that limb strength among DMD boys was 40-60% of healthy controls at an age of 6 years and only 10-33% of healthy controls at the age of 11 years.

Second, timed motor performance tests, interpreted as velocity measures, are reliable and feasible in the majority of the population from age 4 and older. While these are limited in practice by onset of functional disability, inclusion of zero-value velocities can account for those who cannot perform the tests. The tests best demonstrate differences in steroid-treated and naive populations from the youngest ages into the mid-teens, and are most sensitive to 1-year changes between 7–9 and 10–12 years of age for standing from supine, and 7–9, 10–12 and 13–18 years for stair climb and 10-m run/walk. Variability is similar to that for quantitative strength measures for individuals under 10 years of age, and may increase over time as only the stronger individuals in a cohort retain the ability to perform the tests. Vignos lower extremity and Brooke upper extremity functional scales are reliable and feasible across the entire survivable age range of the disease, and demonstrate differences between steroid-treated and naive subgroups. Scores on both measures typically change by one grade level or less over a year period, and are most likely to show functional decline between 7–12 years of age for Vignos and 7–18 years of age for Brooke.

Third, pulmonary function measures are reliable and feasible in nearly all patients over 6 years of age. Although FVC showed significant 1-year increases in the 7–12 year age groups and no significant change in the 13–18 year age group, the percent predicted FVC was significantly decreased in the 7–12 and 13–18 year age groups. These changes indicate the confounding effect of growth in these subjects.

A detailed description of the data across the entire study cohort has been submitted for publication and will soon be available. Analysis of these data suggest that some of these outcome measures are valid for use in clinical trials. However, there is a need to develop sensitive, feasible and reliable assessment techniques that maintain continuity of measurement across as much as possible of the survivable age range of the disease. This includes toddlers who are mostly keeping pace with their peers to non-ambulatory adults who are heavily reliant upon caregivers and adaptive technologies to participate in daily activities. Furthermore, for assessments to be clinically relevant, they should be correlated with functional 'milestones' that are relevant to the subjective experience of the patients themselves.

Use of timed motor performance to predict loss of functional milestones

Our data indicates that milestones of disease progression, such as loss of ability to rise from floor, climb stairs, rise from a chair, ambulate 10 m and self-feed occur in a predictable order, and that loss of those abilities can be predicted by timed functional evaluations. As we previously described in a steroid-naive cohort [8], one application of traditional clinical trial outcome measures to evaluation of clinically relevant functional 'milestones' is the use of the 10-m run/walk to predict time to loss of ambulation. In that study, a timed 10-m walk test of more than 12 s was predictive of loss of ambulation within the next 12 months (100%; p < 0.0001), while individuals who were able to complete the task in less than 6 s were significantly more likely to retain ambulatory ability (100%; p < 0.0001). If properly applied, such predictive statistics could be applied as a clinical trial eligibility criteria and/or outcome measure that are both capable of showing differences between groups over the course of a 1-year time period and are clinically meaningful from the standpoint of patient management.



Figure 2. 10-m run/walk velocities by age using the same data as Figure 1, showing a slope of functional decline that is representative of the population studied, inclusive of non-ambulatory individuals, and that enables comparisons with healthy controls.

North Star Clinical Network for pediatric neuromuscular disease

F Muntoni, A Manzur, E Scott, M Eagle, A Mayhew This UK clinical network was initiated in 2003 to promote best practice in the management of DMD and to provide longitudinal natural history data for this condition. A total of 20 specialist pediatric neuromuscular centers from the UK participate in the network supported by the Muscular Dystrophy Campaign. Standardized assessment protocols and a detailed test manual were developed for ambulant children with DMD via a process of consensus meetings and review. Information from the clinical assessment protocols is entered into the North Star database (online since 2006) by the collaborating centers.

More than 300 subjects were registered on the database at the time the data were collated for the meeting (data accrual continues and the number of subjects on the North Star database currently stands at 500 boys), with a large quantity of longitudinal data available. The database currently holds information from ambulant boys; non-ambulant developments are in progress. The majority of these children are steroid-treated. Data that are systematically collected include the following: North Star Ambulatory Assessment (NSAA, functional scale); timed 10-m walk/run, timed rise from supine; FVC (percent predicted for height and actual); manual muscle testing (percentage MRC2); joint range of movement; functional indicators; genetic mutation and dystrophin status on muscle biopsy; steroid medication (dose, changes, side effects and tolerance); cardiac function, bone health, including dual energy x-ray absorptiometry scan; general health measures, including height, weight and blood pressure; and age at loss of ambulation. A national audit 'Benefits and adverse effects of Glucocorticoids in boys with Duchenne muscular dystrophy: a UK perspective' is currently underway, led by the Dubowitz Centre.

Data from 240 subjects aged 4-12 years were presented at the meeting. Assessments are generally undertaken prior to staring steroids and then at 6-monthly intervals. These data are summarized in Figures 3 & 4. However, the number of subjects in the 11–12 year-old age group is currently small and therefore, this group is not included since it may not be representative. The younger end of the age range (<6 years) may also need to be viewed with some caution, as numbers here are again relatively small. However, these data are presented as they show a trend for improvement in physical abilities as indicated in Figures 3-5. For the North Star dataset, timed rise from supine, timed 10-m 'walk' and the NSAA total score appeared to indicate a greater degree of change in abilities for patients in this age group than other measures used.

Mean FVC percentage predicted for height remained relatively stable in the age range presented, showing a variation of only 2 percentage points in 6–10 year olds. Mean % mean muscle score results showed only a 6 percentage point change across the same age range, and will need further evaluation. Standard deviation for all outcome measures was substantial and showed a general increase with age, particularly from approximately 9 years of age onwards.

The NSAA has undergone further rigorous psychometric review using modern psychometric techniques, in this case Rasch analysis methods [9.10]. Traditional

psychometric methods are those most commonly used for examining scale reliability and validity and are well understood by clinicians and clinical researchers. They are important in our understanding of the measurement basis of rating scales. However they do have significant limitations and Rasch analysis is a measurement diagnostic tool that enables us to comprehensively evaluate the mechanics of scale performance. This modern psychometric method allows examination of a scale's clinical cohesiveness, independence and stability of items. It also informs how response options are performing and identifies anomalous data, which either identifies 'problem' subjects, who misfit the measure or problem items, which misfit the construct. When the NSAA was evaluated using this method, it was shown that the construct of ambulatory function was clearly defined by the 17 items and discriminated between boys ability. There were some issues with misfit (supine head lifts), the breadth of the scale and the inclusion of bilateral items; however, all of these are minor issues and can be addressed. In summary, the NSAA was shown to be a precise and clinically relevant tool [10]. Further analysis is planned on a longitudinal dataset to assess the measure's testretest reproducibility. Should the NSAA continue to meet the criteria necessary under Rasch methodology, conversion of the NSAA from a nonparametric to a parametric measurement tool will be undertaken and further work will be done by the network on correlation with other outcome measures using parametric statistical analysis techniques.







Figure 4. Mean timed 10-m walk/run by age.

North Star: Italian dataset E Mercuri

An Italian clinical network was initiated in 2006 to promote collection of longitudinal natural history data in ambulant children with DMD. The network includes all the 13 leading neuromuscular pediatric centers in Italy. As part of this network, all the centers were asked to participate in training sessions [11] and to use standardized assessment protocols including the NSAA, timed items and the 6MWT at 6 month intervals at least.

Reliability, feasibility & variability of outcome measures

Before collecting natural history data, all the participants were trained in the methods. In order to investigate the level of training needed for achieving a good inter-observer reliability in a multicenter setting, two training sessions were held. Following the first training session there were no difficulties in performing the items and in obtaining adequate video recordings of the session; but the inter-observer reliability was rather



Figure 5. Mean North Star Ambulatory Assessment total score by age.



Figure 6. North Star Ambulatory Assessment score by age.

poor (<0.5). This dramatically improved after a second training session with review and discussion of the videos previously scored (0.995). The level of agreement was maintained even when some videos were rescored after a month, showing high intra-observer reliability (0.95).

Data from 112 subjects aged 4-17 years in the Italian cohort were presented at the meeting in Washington, DC, USA. A focus of the presentation was to understand the correlation between the NSAA and other measures, such as the 6MWT and timed items. Cross sectional data collected at baseline showed that the NSAA had a moderate-to-good correlation with 6MWT and with timed rising from floor; but less with the 10-m timed walk/run test. The 6MWT had better correlation with 10-m timed walk/run test than with timed rising from floor. Some of these data have already been published [12] but further analysis of the same cohort at baseline showed that NSAA and 6MWT performances increased with age up to 7 years, with a clear point of slope change at approximately 7 years (Figures 6 & 7). After the age of 7 years, there was a variable decline in both NSAA scores and 6MWT that appeared to be related to steroid treatment.



Preliminary data on the change observed in these measures in the same cohort over a 12-month period were also shown. One of the advantages of these data is that it has been collected in the last 2 years, therefore reflecting recent standards of care that were shared by all the participating centers. A detailed description of the 12-month data in the study cohort has been subsequently submitted for publication and will soon be available. The results confirm the point of slope change at approximately the age of 7 years for NSAA and 6MWT and define the variability of the change over 12 months. Part of this variability can be explained by different steroid regimens above the age of 7 years, and more consistently above the age of 10 years. Patients treated with daily steroids were overall more stable, compared with untreated patients or those treated with intermittent regimens. These data, therefore, suggest that age and steroid treatment should be carefully considered at the time of designing a clinical trial when deciding stratification criteria. The data on variability over time and in specific age groups will also be useful when calculating sample size.

The Motor Function Measure dataset C Payan, C Berard

The motor function measure (MFM), consisting of 32 items (tests) has been designed for neuromuscular disorders and is applicable to patients at various stages of functionality – ambulant or not, mildly severe to very severe. The first phase validation study, which included analysis of 303 patients, aged 6–60 years, with various disorders (72 were DMD patients) was published in 2005 [13], including details concerning reliability and construct validity. Principal component analysis demonstrated three independent dimensions:

- D1 standing and transfers (13 tems);
- D2 axial and proximal motor function (12 items);
- D3 distal motor function (7 items).

Summary sub-scores and a total score are calculated, expressed as a percentage of maximum possible value. Responsiveness has been studied in a subgroup of 152 patients of which 41 were DMD patients. Data concerning the DMD patients have been published separately [14].

The data presented at the meeting included DMD patients who were not treated with steroids. In this population, total score of the MFM was related to age (Figure 8), with a rapid decline between 5 and 15 years. All patients with scores below 50% were non-ambulant. The progression of the total score is represented in Figure 9 for 41 patients evaluated twice at a mean of 16-month intervals (range 11–22 months). Differences

in scores were transformed by linear interpolation to obtain annual score change. Overall responsiveness expressed with standardized response mean (SRM) was large for the total score (0.86). In ambulant patients, the SRM of the total score and D1 were large (1.32 and 1.22, respectively) while in non-ambulant patients, the SRM were smaller but still substantial (Table 2). The analysis of six cases that lost ambulation between the two assessments showed that a D1 score of 40% was predictive of loss of ambulation 1 year later. A case-control comparison of 12 patients treated 1 year with steroids versus age-matched nontreated patients showed a stability of the MFM scores in the treated group, with an effect size of 1.04 for the total score and 1.22 for the D2 score. At present, a reduced 20-item



Figure 8. Total score of Motor Function Measure-32 versus age in 72 Duchenne muscular dystrophy patients.

version of the MFM is undergoing validation in children below 6 years of age and has shown psychometric properties similar to those of the full version (publication in preparation for the reliability and construct validity). A database has been designed to continue collection of data from many clinical sites in France and abroad. This will enable confirmation of preliminary results in larger samples (over 150 DMD patients available).

The Danish dataset B Steffensen

The Danish dataset consists of information from children and young men with DMD who are registered with

the Danish National Rehabilitation Centre for Neuromuscular Diseases, including 160 individuals from the age of 3-48 years. Half of the participants are 20 years old or older. The incidence of new patients registered is six per year and the number of individuals registered indicates that this represents the total DMD population in Denmark. Diagnosis is based on clinical and genetic criteria used internationally and (to differentiate between Becker Muscular Dystrophy [BMD] and DMD) a biopsy is taken in the early stage to assess the amount of dystrophin. Before genetic testing was possible (in the 1980s), the criteria

described by the European Neuromuscular Centre were used. To differentiate between Becker and DMD the clinical criteria for DMD was, as described by European Neuromuscular Centre, that independent ambulation was terminated before the 13th birthday. Since 1995, participants have been assessed regularly once a year in the Danish National Rehabilitation Centre for Neuromuscular Diseases up till the age of 18 years. After the age of 18, they are seen at longer intervals. The assessments are performed in accordance with the study protocol used in 1995 [15] for ambulatory boys and later the Egen-Klassification (EK) scale became standard for non-ambulatory boys. Reliability was assessed early and repeated when a new physiotherapist was introduced.





Meeting Proceedings Bushby & Connor

Table 2. Differences in motor function measurescores (between baseline and second visit)expressed as annual rate of change (mean ± SD)and standardized response mean.

	Mean change ± SD	SRM		
Total group (n = 41)				
Total score	-5.8 ± 6.3	0.86		
D1	-4.9 ± 10.4	0.47		
D2	-7.7 ± 11.2	0.68		
Subgroup ambulant patients at first visit (n = 11)				
Total score	-7.9 ± 6.0	1.32		
D1	-17.2 ± 14.1	1.22		
Subgroup non-ambulant patients at first visit (n = 30)				
Total score	-5.0 ± 6.3	0.79		
D2	-9.4 ± 12.5	0.75		
Subgroup age >14 years at first visit (n = 16)				
Total score	-3.9 ± 6.4	0.61		
D2	-3.9 ± 8.1	0.48		
D3	-10.8 ± 18.5	0.58		
SRM: Standardized response	mean.			

Assessments consist of manual muscle test (40 muscle groups), hand held myometry (14 muscle groups), range of motion (20 joints including jaws), FVC and FVC%. Functional tests for ambulatory patients include: Hammersmith motor ability; timed tests: 10-m walk, standing from lying supine, four stairs up and down. Functional test for non-ambulant patients is the EK scale. When FVC is reduced below 50% of reference values respiratory assessments are supplemented with sleep studies once a year, to assess hypercapnia during the night.

Treatment with steroids was gradually started in the beginning of 2000 and it was not until 2005 that all boys were treated with steroids. This means that the cohort includes a population of patients born between 1983 and 1990 who were never treated with steroids but were examined prospectively and regularly. The purpose of this was to be able to describe the natural history later on when all patients would be on steroids. This

Table 3. Annual change in select outcome measures of 19non-ambulatory boys.				
Variables	Mean	Minimum	Maximum	
EK sum (score)	1.3	0.0	3.8	
MRC%	-2.0	0.7	-7.9	
FVC%	-6.3	-1.3	-16.9	
EK: Egen-Klassification scale; FVC: Forced vital capacity MRC: Medical Research Council				

EK: Egen-Klassification scale; FVC: Forced vital capacity MRC: Medical Research Council manual muscle test. steroid-naive group (n = 46) was presented at the meeting with respect to their age at clinical events and assessments with EK and FVC in relation to clinical events after loss of ambulation. Data were presented from all patients, alive as well as those who had died during the period of follow-up. A total of 15 of the 46 individuals were not alive in 2010.

Clinical events in relation to age

In the Danish cohort the mean age at loss of ambulation was 9.1 years and the mean age at scoliosis surgery, if done, was 14 years (nine out of 46 did not develop severe scoliosis and did not have surgery). The mean age for the development of hypercapnia requiring night time-assisted ventilation was 17 years, and the mean age when day and night hypercapnia required day and night-assisted ventilation among those patients with hypercapnia was 20 years. For the 15 who died, the average age at death was 18 years. Sensitivity of EK sum score, FVC% and manual muscle test (MRC%) was tested in a previous longitudinal study [16] of 19 nonambulatory boys with DMD as annual change and showed mean, minimal and maximal change per year, as seen in Table 3.

EK & FVC% as predictors of hypercapnia day & night In the older age group of patients a key milestone of disease progression is the requirement for nocturnal ventilator support. The need for assisted ventilation during the day and night due to hypercapnia and hypoxia was mainly predicted by the EK sum score and FVC% in this group, as shown in an earlier study [17]. An EK sum score >20 and FVC% <30 indicated a need of assisted ventilation up to 18 months after assessment. Neither an EK sum >20 nor FVC% <30 alone was sufficient to predict the need of assisted ventilation. This dataset will be further examined to determine the effect of steroid treatment, as this is a group of patients already followed prospectively who are in adulthood. This population has the potential to provide a very good dataset on useful outcome measures for older boys.

Clinical evaluation data in the United Dystrophinopathy Project KM Flanigan for the United Dystrophinopathy Project Investigators

The United Dystrophinopathy Project (UDP) is an NIH-funded consortium of seven centers studying genotype-phenotype characteristics and natural history in patients with the dystrophinopathies (DMD, BMD, and intermediate muscular dystrophy [IMD]). Subjects have historical data extracted from their records, with particular attention to age at symptom onset, cardiac data, and medication data (in particular

the history of steroid use). Each undergoes a standardized examination consisting of manual motor testing (modified MRC [mMRC] scale); timed functional testing (time to climb four stairs, arise from floor, transit 10 m); Brooke and Vignos limb scales; handheld myometry of elbow flexion and knee extension; goniometry; and pulmonary function testing (including FVC and FEV₁). These values are assessed at yearly visits whenever possible.

Mutational analysis is performed (or confirmed) at the Utah Genome Center (RB Weiss, Director), using a combination of exon copy number testing and a direct sequencing assays developed at Utah [18]. Enrollment is open to any patient with:

- Clinical features of DMD/BMD and an X-linked family history;
- Muscle biopsy demonstrating absent or altered dystrophin expression (immunofluorescence, immunohistochemistry, immunoblot);
- DMD gene mutation in previous clinical testing.

Genotyping

A major emphasis of this project has been on genotyping, both to further define the mutational spectrum, and to identify molecular mechanisms associated with milder phenotypes than predicted from genomic mutational analysis. Toward that end, the UDP and the Utah Genome Center have published the mutational spectrum in 1111 dystrophinopathy patients, of whom 891 had clearly defined phenotypes [19]. This paper addressed SNP distribution at the locus, and raised the hypothesis that nonsynonymous SNP may be related to severity of symptoms. Other genotypically oriented reports making use of UDP data addressed the distribution and mechanisms of exon duplications [20]; characterized the mutational distribution in an unbiased clinical survey [21]; described the presence of pseudoexon mutations as an under-recognized mutational class [22]; defined the first founder allele in the DMD gene [23] and experimentally defined the mechanism by which this nonsense mutation was associated with a very mild phenotype [24]; described a lack of association between nonsense

mutation read through and baseline phenotype [25]; described correlations between deletion location and cardiomyopathy phenotype in BMD patients [26]; defined the clinical spectrum in a cohort of manifesting carries of dystrophinopathy [27]; and characterized splice site metrics and their relationship to nonsense mutations in the setting of nonsense-associated BMD [28]. Current enrollment is listed in Table 1; at the time of the meeting, 799 dystrophinopathy patients (including DMD, IMD and BMD) had undergone at least one evaluation. Patients with subsequent yearly visits included 462 (year 2), 260 (year 3), 128 (year 4) and 42 (year 5).Data from an interim analysis of the first 500 enrollees were presented. Of these, 371 had DMD, of whom 174 were steroid-naive at enrollment, with a mean age of 10.5 ± 5.9 years; 32 had IMD, with a mean age of 14.0 ± 6.8 years; and 97 had BMD, with a mean age of 24.2 ± 16.0 years. The distribution of steroid use versus no steroid use was equal in the DMD group (Table 4).

Two examples of analysis among the first 500 subjects

Modified Medical Research Council scores

Although use of the mMRC in DMD trials has repeatedly been found useful, dating to the CIDD trials, in multicenter trials inter-rater reliability can be challenging. Testing among evaluators at the seven UDP centers provided data on inter-rater correlation coefficients for each mMRC test. This allowed analysis of a subset of tests with International Co-ordinating Committee for Spinal Muscular Atrophy of >0.7 on two testing dates. Among these, four muscle functions (shoulder abduction, elbow flexion, elbow extension, and knee extension) could be used to generate a 'mini-mMRC' score, which correlated highly (p < 0.001) with the full mMRC battery (consisting of 18 muscle groups). This suggests that a much briefer mMRC battery may be substitutable in clinical trials, although further validation is required.

Steroid use

Among those boys with DMD who had lost ambulation, those on any steroid regimen for >6 months walked significantly longer (median age at loss of ambulation = 12.0 years) than those on any regimen for <6 months (median = 10.0 years) (Figure 10). There was no significant difference between age at loss of ambulation and whether the patient was on a daily or nondaily regimen (p = 0.77).

Table 4. United Dystrophinopathy Project dataset: steroid usage cross type of muscular dystrophy.					
Steroid use	DMD N (%)	IMD N (%)	BMD N (%)		
\geq 6 months	168 (45.3)	14 (43.8)	17 (17.5)		
Between 1 and 6 months	17 (4.6)	0 (0)	2 (2.1)		
Never or ≤1 month	174 (46.9)	17 (53.1)	76 (78.4)		
Not interpretable	12 (3.2)	1 (3.1)	2 (2.1)		
BMD: Becker muscular dystrophy; DMD: Duchenne muscular dystrophy; IMD: Intermediate muscular dystrophy.					

Meeting Proceedings Bushby & Connor



Figure 10. Proportion not using a wheelchair fulltime in Duchenne muscular dystrophy patients by age.

At present, enrollment in the UDP has closed, and data are being cleaned for a detailed analysis of all the historical phenotypic and prospective clinical examination data. Although further genotype-phenotype studies are anticipated, the current emphasis is on analyzing this extensive dataset, and publications within the near term are planned. Meanwhile, the database has played a critical role in clinical trial readiness, having been used to identify patients for past and ongoing studies of nonsense mutation read-through; exon skipping; myostatin inhibition; and more focused outcome measure studies (as a preparation for gene transfer trials).

The 6MWT in Duchenne/Becker muscular dystrophy clinical trials: experience from the international study of Ataluren (PTC124[®]) in nonsense mutation Duchenne/Becker muscular dystrophy

A Reha, GL Elfring, L Atkinson, J Barth

Ataluren is the first investigational drug designed to promote read-through of a premature stop codon in the mRNA, resulting from a nonsense mutation in the DNA, leading to the production of a full-length, functional protein. A Phase IIb, randomized, double blind, placebo-controlled, dose-ranging study was conducted to assess the efficacy and safety of ataluren in males \geq 5 years with nonsense mutation Duchenne/Becker muscular dystrophy documented by dystrophin gene sequencing. Enrolled patients were ambulatory and were stratified by age, corticosteroid use, and baseline 6-min walk distance (6MWD). In this study, patients received high-dose ataluren, low-dose ataluren or placebo orally for 48 weeks. Outcome measures included 6MWD, muscle function and strength, and safety.

The primary end point of this Phase IIb study was change in 6MWD. The 6MWT is a global measure of endurance and muscle function that has been used to evaluate functional capacity in neuromuscular diseases and has served as the basis for marketing approval of several drugs. In a short-term observational study of 6MWT performance in 21 ambulatory boys with DMD and 34 typically developing boys, McDonald *et al.* reported that a modified 6MWT is feasible, safe, and reliable in boys with Duchenne/Becker muscular dystrophy (DBMD) who have not yet transitioned to fulltime wheelchair use [29]. Based on these observations, the 6MWT was adopted as the primary outcome measure in this Phase 2b study.

This Phase IIb study enrolled 174 subjects (median [range] age = 8 years [5-20]) at 37 sites in 11 countries. Given the size and scope of this trial, a number of efforts were made in advance to standardize the conduct of the 6MWT and other outcome measures across the participating study sites. These efforts included the formation of a Clinical Evaluator Training Group, standardization of equipment used for the outcome measures across multiple sites, development of a detailed Clinical Evaluator Training Manual and related training materials (e.g., worksheets and training videos), identification of at least two qualified physiotherapists at each participating site, centralized training and re-training after 1 year, assessment of pretreatment data, and provision of ongoing feedback to the study sites during the conduct of the study.

Over the course of the study, the standard deviation of the change in 6MWD (a measure of variability) increased to approximately 90 m by week 48, providing evidence of the heterogeneous nature of disease progression in DBMD over the course of approximately 1 year. Greater variability, as determined by a comparison of the coefficient of variation values, was also observed for specific outcome measures of muscle function and strength. Advantages of the 6MWT relative to other outcome measures include a more global measure of patient functional ability, a direct assessment of clinically meaningful information, and the precedence of acceptance by regulatory authorities.

Experience gained from this study indicates that the 6MWT is a valid outcome measure of global patient function and should remain a key end point in studies of ambulatory DBMD patients. Furthermore, the change in variability during the 48-week measurement observed

in this study should be considered when deciding the sample size of future studies.

The German MD-Net DMD-cyclosporin A trial R Korinthenberg, J Kirschner, J Vry

The German MD-Net DMD-cyclosporin A trial represents one of the largest controlled trials carried out to date in DMD. A total of 153 patients were randomized in 11 centers during 3 years, 76 to placebo and 77 to cyclosporin A (CsA). The inclusion criteria were at least 6 years of age, ability to walk independently for at least 50 m, diagnosis of DMD proven by clinical symptoms and genetic testing and/or muscle biopsy analysis and ability to participate in the evaluation. Key exclusion criteria included: previous treatment with glucocorticoids and new or changed treatment with food supplements during 3 months prior to inclusion. A total of 73 patients were <7 years of age and 72 were >7 years. During the first 3 months patients were treated with placebo or CsA. From month 3 through to month 15, both groups additionally received intermittent prednisone (0.75 mg/kg, 10 days on/10 days off). The outcome measures included: MRC% (28 muscle groups); handheld Citec-myometry (5 muscle groups, dominant side, best of three trials); ability and time to stand up from supine; Gowers Score and ability and time to walk 10 m. In addition to investigator instructions in writing, photographs and video, central investigator training was repeated every year [30]. As no difference between the placebo and treated groups were found, the following correlations were calculated in the whole group of patients.

MRC% and myometry test-retest reliability for time 0 and 3 months was good (>0.7) for all muscle groups combined and for the individual muscle groups. The same is true for the timed function measures (>0.83). However, both MRC% and myometry results depended significantly on trial sites, which was not the case for the timed function-tests. Loss of the ability to stand up from floor during the following 15 months could be significantly predicted by a MRC% total <60% or a total myometry strength <180 N at start of the study (MRC% 89 vs 44% of patients, myometry 81 vs 55% of patients) MRC% total and MRC% for shoulder, elbow, hip and knee discriminated significantly between age groups greater and less than 7 years of age, but they did not change during treatment course (0-3 months, 0-15 months). Hand-held myometry of knee extension discriminated significantly between age groups, too. This was not the case for hip flexion and shoulder abduction, and inconsistently for elbow flexion. Correlation between MRC% and myometry was moderate for most muscle groups. A correlation >0.6 was found for total score, elbow flexion and knee-extension.

Timed tests (in those being able to perform)

discriminated significantly between age groups, but they did not change over time. At the 15-months investigation, we found correlations >0.6 between time to walk 10 m with MRC% total, MRC% knee, MRC% elbow and myometry knee extension, but not for the other measures and not for Gowers time. However, Gowers score correlated >0.6 with MRC% total, MRC% hip, MRC% knee and myometry knee extension.

In summary, test–retest reliability was good for all measures, although MRC% and myometry showed an investigator effect. Both MRC% and myometry were correlated to the timed function test and Gowers score, and were able to predict loss of the ability to rise from floor. This large study demonstrates the ability to standardize assessment of children with DMD across a large number of centers. In addition, the data collected indicate that there can be value in the outcome measures used in predicting clinically meaningful end points to the patient and family. Such measures could be discussed as clinical surrogate measures for physical ability and quality of life.

Outcome of facilitated discussion Setting the foundation of our understanding of the natural history of DMD

Historical data on the natural history of disease in DMD are available from multiple international studies and publications. These data show that the progression of DMD has a predictable trajectory. Early on, boys have a period of gaining motor skills, albeit at a slower level compared with their healthy peers, and thus do not reach comparable physical performance. A plateau phase is followed by the progressive loss of motor function including the ability to rise from the floor, climb stairs, walk independently, self-feed, and sustain adequate nocturnal ventilation without assistance. In nonsteroid treated patients these phases can be mapped very specifically and are reflected in the MFM data presented here and the Danish dataset, both of which are predominantly reporting data from steroid naive patients. Alongside physiotherapy and general medical care, corticosteroid treatment has had the major positive impact to date on loss of motor functional ability (specifically loss of ability to get up from the floor, climb stairs and walk independently). Following loss of ambulation, glucocorticoid treatment also significantly prolongs the ability to self-feed and reduces the requirement for scoliosis surgery. Data on long-term use of glucocorticoids are beginning to show that they also improve respiratory function, resulting in a delay in the mean age at requirement for assisted ventilation. The 'new' natural history of DMD can, therefore, also be mapped with regard to preserved function in each of these parameters. Based on the available data and historical studies of DMD, the group arrived at a consensus about the overall contemporaneous milestones in DMD in the steroid era as summarized in Figure 11.

This framework and the data from the natural history studies can be used to inform the design of clinical trials of new candidate therapeutics for DMD. To be able to observe differences in decline of function, inclusion of patients between 6 and 10 years of age appears optimal. Predictive data from the CINRG dataset can be utilized to decrease variability and reduce sample size by identifying boys with a trajectory of decline that is likely to be observed in a reasonable time period (e.g., 6-12 months). This will need to be balanced against consideration of excluding patients who may have such an inevitable pace of progression that improvement with a study drug may represent a high bar. Studies designed to provide evidence for efficacy of a new treatment would need to address the effect of glucocorticoids on disease progression, both as entry criteria and as concomitant medications during the period of observation.

An ongoing task for the working group of investigators and their natural history datasets is the correlation (and determination of predictive value) of the various outcome measures to those functional milestones with real clinical relevance to the patient and family, such as loss of ability to get up from the floor, climb stairs, walk independently, self-feed and sustain adequate overnight ventilation without support. These intermediate 'patient-oriented' milestones in disease progression have been neglected to date, in deference to the obvious major milestone of loss of ambulation. However, trials predicated on the time to loss of ambulation are likely to be too long to be practical. The





kind of work presented by the CINRG group correlating functional tests with age at loss of ambulation can now be extended to study these intermediate but still highly clinically relevant timepoints in the progression of the disease.

Another important issue in therapeutic trials is variability. Inevitably clinical trials will be limited by the availability of patients given the orphan nature of DMD and the subpopulations that will be determined by entry criteria. Within the working group's evaluation, we will need to understand more about the variability of the outcome measurements studied. Data from the PTC therapeutics trial control dataset and others suggested that variability increased as severity of disease increased – as boys approached the time of losing the ability to walk altogether, their performance on a specific task was much more variable. This will have impact on the stratification and design of trials.

Filling in the gaps

Analysis of the datasets collected as part of the working group meeting indicates that an impressive amount of contemporaneous natural history data exists on the progression of DMD in a large number of children using a range of relevant outcome measures. All studies are multicentric and confirm the feasibility and reliability of repeated evaluations performed at multiple sites. These datasets are enriched for the ambulant DMD population over the age of 6 years that appear optimal for inclusion in pivotal trials designed to evaluate the treatment effect of new therapeutics. Nevertheless, it will also be important to assure that other populations of DMD boys are included in the label for any new therapy. It is anticipated that both younger ambu-

> lant boys and older non-ambulant patients will be included in other pre- or post-approval clinical trials and studies and in registries. Very few children under the age of 4 years have been enrolled in any systematic natural history study, a situation being addressed by the five sites within the Muscular Dystrophy Association (MDA) DMD Clinical Centers of Excellence Network (Boston Children's Hospital, Boston MA, Nationwide Children's Hospital, Columbus OH, University of Minnesota, MN, University of California at Davis and Washington University in St. Louis, MO, USA, as well as Newcastle upon Tyne, UK) to address the gaps in our knowledge in this age range. While

the mean age of diagnosis in DMD remains approximately at 4 years old, this is a scarce group of children relatively infrequently diagnosed. Intensive study is needed to ensure that these patients are identified and assessed systematically.

The goals of the DMD infant and young child (aged 1 month to 5 years) study are to:

- Establish fine and gross motor development as an effective outcome measure in infants and young children;
- Assess language and cognitive development in infants and young children;
- Assess ultrasound of biceps and quadriceps as a marker of disease progression;
- Assess the burden for the primary care giver of DMD infants and young children.

Evaluations performed in this study include height, weight, vital signs, ultrasound (at selected sites), the Bayley Scales of Infant and Toddler Development (including cognitive, receptive and expressive language, fine and gross motor, social, emotional and adaptive behavior), the NSAA, the Hammersmith Functional Mobility Scores Extended, and a caregiver burden scale.

In addition to the Danish study reported as part of the working group, the MDA centers are undertaking a study in older boys/men with DMD who are non-ambulatory. The goals of the latter study are to:

- Establish optimal and reliable clinical assessments in this population;
- Evaluate patient reported outcomes using the Individualized Neuromuscular Quality of life Questionnaire (INQoL) in adults with DMD and;
- Test caregiver burden in the primary caregiver. Evaluations include vitals, height, weight, hand-held dynamometry, manual muscle testing, active and passive range of motion, the Jebsen, the nine hole peg, pulmonary function, the INQoLs (in individuals 18 years and older) and a caregiver burden scale.

Further studies of the non-ambulant DMD population are under way in France.

Better understanding of the natural history and outcome measure performance in these patients will be important in clinical development of new DMDspecific therapeutics. These additional studies will extend our knowledge of the functional abilities of boys with DMD into new age ranges and allow us to identify appropriate measures to monitor potential therapeutics in these young children and young men.

Future perspective

As new and innovative DMD-specific investigational agents move into clinical trials it is critical to consolidate our knowledge and delineate the corpus of the natural history data available to the scientific community. The purpose of this workshop was to bring together leading investigators from around the world to share available natural history data and to launch a collaborative effort designed to support the development of new DMD therapeutics. The prospective studies and retrospective analyses presented at the meeting represent a major body of work and a strong foundation and important resource for planning and interpreting clinical trials.

The studies presented here tell a strong and consistent story about the current natural history of DMD and, for future trials, we have an opportunity to map the outcome measures used in trials against this natural history. While relatively under-represented in this dataset to date, the 6MWT is, nonetheless, showing correlation with intermediate milestones of progression in other studies. The 6MWT is being utilized in clinical trials of antisense oligonucleotide treatment currently and will likely represent a mainstay outcome measure. Work by investigators in the CINRG network provide useful approaches to the interpretation of 6MWT data that can be used in patient selection and data analysis. As the body of data on 6MWT accumulates (e.g., with the additional data from PTC and others) the collaborative dataset will serve as a robust 'laboratory' for addressing issues relevant to subsequent trials. In addition, the collaborative dataset can be utilized to define the relationships among clinical outcomes that will be useful in selecting alternatives to timed tests similar to the 6MWT.

Functional tests that 'matter' can be better defined on the basis of the data collected in these and other studies. Clinical meaningfulness is a key point and can be seen in terms of intermediate milestones of disease from the data presented in this meeting. For example, a 30 m difference in the 6MWD was correlated with patientrelated outcomes in other populations. Changes in the various outcome measures discussed in the meeting are amenable for application to trials and simulations of changes in end points with treatment may be feasible using steroid data. Children with DMD started on steroids (depending on their stage of the disease) will show improvement in the vast majority of the parameters collected in these studies, thereby providing a degree of 'positive control' for treatment effect.

A robust natural history dataset is important to define disease stage in DMD. Staging of the disease and description of the impact of changes in stage is important for milestone/event analysis in trials. Investigators at the meeting identified a number of initiatives in which collaborative datasets could be used to assess the validity of disease staging and to simulate outcomes in a 'virtual' control group of DMD boys. Innovative approaches to trial design need to be explored in DMD therapeutics (e.g., adaptive trials) and these will be facilitated by the ongoing work of the group.

It was an enormously strong precedent to have PTC at the table as the first drug company to have completed a regulatory directed study in DMD. This group can act as a catalyst to drive acceptable practice with industry including the iterative study of placebo data. The group believes that it would be highly advantageous to look at baseline data from patients enrolled in trials as soon as possible. On this basis, it would be important to encourage companies to expect that level of involvement. The group assembled and the networks represented at the meeting represent a composite group to interact with industry.

Alongside clinically relevant outcome measures and their further definition, the collaborative effort initiated in this working group can provide other tools that can be adapted and further correlated for their usefulness in trials. A good example is the ongoing parallel effort to standardize the analysis of dystrophin expression where the absence of standardization is a problem. Work is ongoing in several centers to correlate dystrophin expression with clinical disease expression in DMD and BMD. The collaborative nature of the group also provides an advantage in terms of gathering centralized information, as training and flexibility is an issue. There is the need to share many different measures and develop a trial mindset across the community so that adequate numbers of sites are available for the trials that may be needed. Initiatives within TREAT-NMD and with its collaborating groups such as the Registry of Outcome Measures, patient registries and care and trial site registries contribute to this development.

Within clinical trial networks and broader networks such as TREAT-NMD, it is feasible to see developments such as a clinical evaluator academy where a 'toolkit' of trained evaluators can be developed in performing core outcome measures. Although it is clear that the field strives for the 'perfect' outcome measure, in the end the trick is consistency rather than perfection of choice of measure. Measures need to be selected carefully depending on the stage of the disease and show responsiveness to change and correlation to the intermediate clinical end points defined through this collaborative work.

The resources of the working group address the several core issues in drug development for DMD and will be valuable to academics, pharmaceutical sponsors and regulatory agencies. The group, drawn together under the auspices of a collaborative network has the advantage of independence and thus provides an objective contribution to the field. In addition it can provide a means of standardization in trial design and consistency in supportive care and outcome measures that will be important as candidate drugs are evaluated. The data presented at this meeting, where the use of steroids has been the major driver for a change in the natural history in the ambulant DMD population are a key example of this, but care in DMD is complex and multidimensional, and failure to address this complexity will jeopardize trial success. This is an issue which requires consistency as different companies enter the trial arena for DMD.

It was acknowledged that the mission and timeline for companies entering the trial arena is different from that in the academic arena, and that this interplay between different communities requires constant dialogue. With this in mind, it was suggested that this group should reconvene on an annual basis, and that it should be a principle of the meeting that companies engaged in trials in DMD should be asked to participate and/or contribute within a consistent format that will facilitate sharing of nonconfidential data. PTC led the way in the current meeting in their full and frank discussion of the placebo data from their completed trial; all meeting participants agreed that such openness could be a gain to the whole field to share and maximize efficiency.

In the end, patients are central to this discussion and we have an ethical obligation not to put children into trials for nothing. We need to ensure that information obtained from trials (even those that are not successful) is useful to the community at large and informs future work. A well-run negative trial thereby fulfils its obligation to move the field forward and not put children through often a rigorous set of procedures in vain. Working as a collaborative network, the view from a coalition of academia, researchers and advocacy is much more powerful than from one body, such as industry alone, especially in addressing a problem that is solvable. DMD has received more attention and study than other neuromuscular diseases, so it is important that this principle is seen as a paradigm for collaboration in other neuromuscular diseases so that we are in a stronger position earlier for other diseases.

The group committed to a number of follow-on activities, including regular calls, a meeting in a year's time, the extension of shared studies in the older and younger populations in DMD and the extension of data sharing into other currently more exploratory end points such as MRI and cardiac evaluations.

Supplementary data

Supplementary data accompanies this paper and can be found at www.future-science.com/doi/suppl/10.4155/CLI.11.113.

Executive summary
 The International DMD Clinical Outcomes Working Group Following key interactions with international regulatory agencies, a workshop was organized by TREAT-NMD and The Children's National Medical Center to review the current status of age-appropriate clinical outcome measures for clinical trials in DMD. Aims of the meeting were to:
 Map the outcome measures currently used in DMD natural history studies and clinical trials, and highlight and assess the data currently being collected; Establish the current expected clinical course of the disease from these contemporary datasets; Determine what her data quict to define the relationship between the outcome measures and milestones of disease programming.
 Determine whether data exist to define the relationship between the outcome measures and milestones of disease progression, Determine areas where the available data clearly describe the disease and where gaps in data highlight the need for further work.
 Preliminary work Previous to the workshop, eight natural history and clinical trial datasets were identified, representing over 1900 patients who are being, or have been, prospectively followed. Investigators from each participating group were asked to present their data in a format that facilitated mapping disease progression and to focus on measures most commonly collected. Outcome measures were described in terms of their sensitivity, reliability and applicability to clinical trials as well as their relationship to disease progression.
 Meeting proceedings: presentation of the available data Longitudinal assessment in confirmed Duchenne muscular dystrophy: Cooperative International Neuromuscular Research Group Study. North Star Clinical Network for pediatric neuromuscular disease. North Star: Italian dataset.
 The Motor Function Measure dataset. The Danish dataset. Clinical evaluation data in the United Dystrophinopathy Project. The 6MWT in Duchenne/Becker muscular dystrophy clinical trials: experience from the international study of Ataluren (PTC124[®]) in nonsense mutation Duchenne/Becker muscular dystrophy. The German MD-Net DMD-cyclosporin A trial.
Outcome of facilitated discussion From the data presented, the natural history of DMD was shown to have a predictable trajectory that was very consistent
 across datasets. Data from two datasets (MFM and Danish datasets) show that the phases of the disease can be mapped very specifically in steroid naive patients.
Alongside physiotherapy and improve general medical care, glucocorticoid treatment has had the biggest positive impact to date on loss of motor function ability during the ambulant phase, and ability to self-feed and reduce requirement for scoliosis during the non-ambulant phase.
 Data on long-term use of glucocorticoids are beginning to show that they also improve respiratory function. To be able to observe differences in decline of function, inclusion of patients between 6 and 10 years of age appears optimal. Predictive data can be utilized to decrease variability and reduce sample size by identifying boys with a trajectory of decline that is likely to be observed in a reasonable time period (e.g., 6–12 months).
 Studies designed to provide evidence for efficacy of a new treatment need to address the effect of glucocorticoids on disease progression.
The studies presented tell a strong and consistent story about the current natural history of DMD, and for future trials we have an opportunity to map the outcome measures used in trials against this natural history.
 Future perspective Ongoing tasks include: correlation of the various outcome measures to those functional milestones with real clinical relevance to the patient and family; understanding more about the variability of the outcome measurements studied; and a better understanding of the natural history in the very young and non-ambulant patients. While relatively under-represented in the datasets presented, the 6MWT is showing correlation with intermediate milestones of
 progression in other studies. This test is currently being used in exon-skipping trials and data needs to be further accumulated to address its utility as a mainstay outcome measure. As the body of data on the 6MWT grows, the collaborative dataset will serve as a robust 'laboratory' for address issues relevant to future trials.
 Staging of the disease and description of impact of changes in stage is important for milestone/event analysis in trial design. Investigators at the meeting identified a number of initiatives in which collaborative datasets could be used to assess disease staging and simulate outcomes in a 'virtual' control group of DMD boys.

Meeting Proceedings Bushby & Connor

Financial & competing interests disclosure

The Working Group would like acknowledge Cure Duchenne, the Foundation to Eradicate Duchenne and Ryan's Quest for their generous sponsorship of this meeting. This publication was in part supported by Award Number UL1RR031988 from the NIH National Center for Research Resources: its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Center for Research Resources or the NIH. TREAT-NMD is supported through Priority 1 (Life Sciences, Genomics and Biotechnology for Health) of the European Union's FP6 under contract number LSHM-CT-2006-036825. Abby Bronson and Emma Heslop provided vital project management support for this group and ongoing activities. The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

No writing assistance was utilized in the production of this manuscript.

International DMD Clinical Outcomes working group

R Ted Abresch, University of California, CA USA^{‡§}

L Atkinson, PTC Therapeutics, NJ, USA^{+§}

J Auld, King's College London, $\mathsf{UK}^{\!\scriptscriptstyle +\!\!\!+\!\!\!}$

J Barth, PTC Therapeutics, NJ, USA§

C Berard, Hôpital Femme Mère Enfant, Bron, France[§]

A Bronson, Children's National Medical Center, Washington, DC, USA^{{\scriptscriptstyle \#}}

P Clemens, University of Pittsburgh, PA, USA[‡]

A Cnaan, Washington University, Washington, DC, USA⁺

V Cwik, Muscular Dystrophy Association, Tucson, AZ, USA⁺

T Duong, Children's National Medical Center, Washington, DC, USA^{\ddagger}

M Eagle, International Centre for Life, Newcastle Upon Tyne, $\mathsf{UK}^{\scriptscriptstyle 4 \sharp}$

GL Elfring, PTC Therapeutics, NJ, USA§

K Flanigan, The Ohio State University, Columbus, OH, USA for the United Dystrophinopathy Project Consortium $^{\rm ts}$

J Florence, Washington University, MO, USA[‡]

EK Henricson University of California, CA, USA^{‡§}

E Heslop, Newcastle University, Newcastle upon Tyne, UK⁺⁺ EP Hoffman, George Washington University, Washington, DC, USA⁺

A Kennedy, Muscular Dystrophy Association, Washington, DC, USA $^{\scriptscriptstyle \pm}$

J Kirschner, Albert-Ludwigs-University Freiburg, Freiburg im Breisgau, Germany $^{\rm s}$

R Korinthenberg, Albert-Ludwigs-University Freiburg, Freiburg im Breisgau, Germany[§]

RT Leshner, Children's National Medical Center, Washington, DC, USA $^{\scriptscriptstyle \pm}$

AY Manzur, Great Ormond Street Hospital for Children (GOSH), London, UK^{\S}

A Mayhew, International Centre for Life, Newcastle Upon Tyne, $\mathsf{UK}^{\mathtt{15}}$

C McDonald, University of California, CA, USA^{‡§}

E Mercuri, UniversitàCattolica del SacroCuore, Rome, Italy^{++§} F Muntoni, Great Ormond Street Hospital for Children (GOSH), London, UK[§]

C Payan, Hôpital Pitié-Salpêtrière, Paris, France^{‡§}

SW Peltz, PTC Therapeutics, NJ, USA§

A Reha, PTC Therapeutics, NJ, USA^{‡§}

E Scott, Muscular Dystrophy Campaign, London, UK[§]

BF Steffensen, National Rehabilitation Centre of Excellence in Neuromuscular Disorders, Aarhus, Denmark $^{\rm ts}$

J Vry, Albert Ludwigs University Freiburg, Freiburg im Breisgau, Germany[§]

G Walter, University of Florida Cancer Center, FL, USA[‡]

¹TREAT-NMD is a FP6 funded network of excellence, EC grant 036825. www.treat-nmd.eu

*Meeting Participant.

§Section Author.

[®]Editorial Assistance.

Bibliography

- Hoffman EP, Bronson A, Levin AA *et al.* Restoring dystrophin expression in Duchenne muscular dystrophy muscle. *Am. J. Path.* 179(1), 12–22 (2011).
- 2 Muntoni F, on behalf of the meeting Steering Committee and the TREAT-NMD Network. The development of antisense oligonucleotide therapies for Duchenne muscular dystrophy: report on a TREAT-NMD workshop hosted by the European Medicines Agency (EMA), on September 25th 2009. Neuromuscul. Disord. 20(5), 355–362 (2010).
- 3 Brooke MH, Griggs RC, Mendell JR, Fenichel GM, Shumate JB, Pellegrino RJ. Clinical trial in Duchenne dystrophy. I. The

design of the protocol. *Muscle Nerve* 4(3), 186–197 (1981).

- 4 Escolar DM, Henricson EK, Mayhew A *et al.* Clinical evaluator reliability for quantitative and manual muscle testing measures of strength in children. *Muscle Nerve* 24(6), 787–793 (2001).
- 5 Vignos PJJ, Spencer GEJ, Archibald KC. Management of progressive muscular dystrophy of childhood. *JAMA* 184, 89–96 (1963).
- 6 Mayhew, JE, Florence JM, Mayhew TP et al. Reliable surrogate outcome measures in multicenter clinical trials of Duchenne muscular dystrophy. *Muscle Nerve* 35(1), 36–42 (2007).
- 7 Florence JM, Pandya S, King WM *et al.* Clinical trials in Duchenne dystrophy. Standardization and reliability of evaluation procedures. *Phys. Ther.* 64(1), 41–45 (1984).
- 8 McDonald CM, Abresch RT, Carter GT et al. Profiles of neuromuscular diseases: Duchenne muscular dystrophy. Am. J. Phys. Med. Rehabil. 74(Suppl.), S70–S92 (1995).
- 9 Hobart JC, Cano SJ, Zajicek JP, Thompson AJ. Rating scales as outcome measures for clinical trials in neurology: problems, solutions, and recommendations. *Lancet Neurol.* 6, 1094–105 (2007).

Clinical outcome measures for trials in Duchenne muscular dystrophy Meeting Proceedings

- 10 Mayhew A, Cano S, Scott E. Moving towards meaningful measurement: Rasch analysis of North Star Ambulatory Assessment in Duchenne muscular dystrophy. *Dev. Med. Child Neurol.* 53(6) 535–542 (2011).
- Mazzone ES, Messina S, Vasco G *et al.* Reliability of the North Star Ambulatory Assessment in a multicentric setting. *Neuromuscul. Disord.* 19(7), 458–461 (2009).
- 12 Mazzone E, Martinelli D, Berardinelli A et al. North Star Ambulatory Assessment, 6-min walk test and timed items in ambulant boys with Duchenne muscular dystrophy. *Neuromuscul. Disord.* 20(11), 712–716 (2010).
- 13 Berard C, Payan C, Hodgkinson I, Fermanian J, MFM collaborative study group. A motor function measure scale for neuromuscular diseases. Construction and validation study. *Neuromuscul. Disord.* 15(7), 463–470 (2005).
- 14 Vuillerot C, Girardot F, Payan C *et al.* Monitoring changes and predicting loss of ambulation in Duchenne muscular dystrophy with the motor function measure. *Dev. Med. Child Neurol.* 52(1), 60–65 (2010).
- 15 Lyager S, Steffensen B, Juhl B. Indicators of need for mechanical ventilation in Duchenne muscular dystrophy and spinal muscular atrophy. *Chest* 108(3), 779–785 (1995).
- 16 Steffensen BF, Lyager S, Werge B, Rahbek J, Mattsson E. Physical capacity in nonambulatory people with Duchenne muscular dystrophy or spinal muscular atrophy: a longitudinal study. *Dev. Med. Child Neurol.* 44(9), 623–632 (2005).

- 17 Steffensen B, Hyde S, Lyager S, Mattsson E. Validity of the EK scale: a functional assessment of non-ambulatory individuals with Duchenne muscular dystrophy or spinal muscular atrophy. *Physiother. Res. Int.* 6(3), 119–134 (2001).
- 18 Flanigan KM, von Niederhausern A, Dunn DM, Alder J, Mendell JR, Weiss RB. Rapid direct sequence analysis of the dystrophin gene. Am. J. Hum. Genet. 72(4), 931–939 (2003).
- 19 Flanigan KM, Dunn DM, von Niederhausern A *et al.* Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostictechniques to a large cohort. *Hum. Mutat.* 30(12), 1657–1666 (2009).
- 20 White SJ, Aartsma-Rus A, Flanigan KM *et al.* Duplications in the DMD gen*e. Hum. Mutat.* 27(9), 938–945 (2006).
- 21 Dent KM, Dunn DM, von Niederhausern AC et al. Improved moleculardiagnosis of dystrophinopathies in an unselected clinical cohort. Am. J. Med. Genet. 134(3), 295–298 (2005).
- 22 Gurvich OL, Tuohy TM, Howard MT *et al.* DMD pseudoexon mutations: splicing efficiency, phenotype,and potential therapy. *Ann. Neurol.* 63(1), 81–89 (2008).
- 23 Gurvich OL, Maiti B, Weiss RB, Aggarwal G, Howard MT, Flanigan KM. DMD exon 1 truncating point mutations: amelioration of phenotype by alternative translation initiation in exon 6. *Hum. Mutat.* 30(4), 633–640 (2009).

- 24 Flanigan KM, Dunn DM, von Niederhausern A et al. DMD Trp3X nonsense mutation associated with afounder effect in North American families with mild Becker muscular dystrophy. Neuromuscul. Disord. 19(11), 743–748 (2009).
- 25 Howard MT, Anderson CB, Fass U et al. Readthrough of dystrophin stop codon mutations induced by aminoglycosides. Ann. Neurol. 55(3), 422–426 (2004).
- 26 Kaspar RW, Allen HD, Ray WC et al. Analysis of dystrophin deletion mutationspredicts age of cardiomyopathy onset in Becker muscular dystrophy. Circ. Cardiovasc. Genet. 2(6), 544–551 (2009).
- 27 Soltanzadeh P, Friez MJ, Dunn D. Clinical and geneticcharacterization of manifesting carriers of DMD mutations. *Neuromuscul. Disord.* 20(8), 499–504 (2010).
- 28 Flanigan KM, Dunn DM, von Niederhausern A et al. Nonsense mutation-associated Becker muscular dystrophy: interplay between exondefinition and splicing regulatory elements within the DMD gene. Hum. Mutat. 32(3), 299–308 (2011).
- 29 McDonald CM, Henricson EK, Han JJ et al. The 6-min walk test as a new outcome measure in Duchenne muscular dystrophy. *Muscle Nerve* 41(4), 500–510 (2010).
- 30 Kirschner J, Schessl J, Schara U et al. Measuring muscle strength in clinical trials. *Lancet Neurol.* 9(11), 1053–1059 (2010).