

Review

“Of Mice and Measures”: A Project to Improve How We Advance Duchenne Muscular Dystrophy Therapies to the Clinic

First Workshop Report: Examining current
findings and opportunities around the emerging
D2.B10-*Dmd*^{mdx}/J (D2/*mdx*) model in context
of the classic C57BL/10ScSn-*Dmd*^{mdx}/J (B110/*mdx*)

Heather Gordish-Dressman^a, Raffaella Willmann^b, Laura Dalle Pазze^c, Arati Kreibich^c,
Maaike van Putten^d, Ahlke Heydemann^e, Laurent Bogdanik^f, Cathleen Lutz^f, Kay Davies^g,
Alexis R. Demonbruen^h, Dongsheng Duanⁱ, David Elsey^j, So-ichiro Fukada^k,
Mahasweta Girgenrath^l, J. Patrick Gonzalez^m, Miranda D. Groundsⁿ, Andy Nichols^o,
Terry Partridge^a, Marco Passini^p, Francesca Sanarica^q, Frederick J. Schnell^p, Dominic J. Wells^f,
Toshifumi Yokota^s, Courtney S. Young^t, Zhong Zhong^u, Christopher Spurney^a, Melissa Spencer^l,
Annamaria De Luca^q, Kanneboyina Nagaraju^{v,1} and Annemieke Aartsma-Rus^{d,1,*}

^aChildren’s National Medical Center, Washington, DC, USA

^bBiozentrum, University of Basel, Basel, Switzerland

^cCharley’s Fund, New York, NY, USA

^dDepartment of Human Genetics, Leiden University Medical Center, the Netherlands

^eDepartment of Physiology and Biophysics, University of Illinois at Chicago, Chicago, IL, USA

^fThe Jackson Laboratory, Bar Harbor, ME, USA

^gDepartment of Physiology, University of Oxford, Anatomy and Genetics, Oxford, UK

^hCenter for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL, USA

ⁱDepartment of Molecular Microbiology and Immunology, Department of Neurology, Department of Biomedical Sciences and Department of Bioengineering, University of Missouri, Columbia, MO, USA

^jSummit Therapeutics, Abingdon, Oxfordshire, UK

¹Shared last authors.

*Correspondence to: Annemieke Aartsma-Rus, Department of Human Genetics, Leiden University Medical Center, the Netherlands. E-mail: a.m.rus@lumc.nl.

^kLaboratory of Molecular and Cellular Physiology, Graduate School of Pharmaceutical Sciences, Osaka University, Osaka, Japan

^lRare Disease Research Unit, Pfizer, CA, USA

^mSolid Biosciences, Cambridge, MA, USA

ⁿSchool of Human Science, the University of Western Australia, Perth, Australia

^oCatabasis Pharmaceuticals, Cambridge, MA, USA

^pSarepta Therapeutics, Inc. Cambridge, MA, USA

^qDepartment of Pharmacy and Drug Sciences, Unit of Pharmacology, University of Bari “Aldo Moro”, Italy

^rDepartment of Comparative Biomedical Sciences, Royal Veterinary College, London, UK

^sDepartment of Medical Genetics, University of Alberta, Canada

^tDepartment of Neurology, Molecular Biology Institute, Center for Duchenne Muscular Dystrophy at UCLA, University of California, Los Angeles, CA, USA

^uWave Life Sciences, Cambridge, MA, USA

^vSchool of Pharmacy and Pharmaceutical Sciences, Binghamton University, New York, USA

Abstract. A new line of dystrophic *mdx* mice on the DBA/2J (D2) background has emerged as a candidate to study the efficacy of therapeutic approaches for Duchenne muscular dystrophy (DMD). These mice harbor genetic polymorphisms that appear to increase the severity of the dystropathology, with disease modifiers that also occur in DMD patients, making them attractive for efficacy studies and drug development. This workshop aimed at collecting and consolidating available data on the pathological features and the natural history of these new D2/*mdx* mice, for comparison with classic *mdx* mice and controls, and to identify gaps in information and their potential value. The overall aim is to establish guidance on how to best use the D2/*mdx* mouse model in preclinical studies.

Keywords: Duchenne muscular dystrophy, DMD, DBA/2J, *mdx*

BACKGROUND

Duchenne muscular dystrophy (DMD) is a severely progressive muscle-wasting disease, caused by mutations in the *DMD* gene, which abolish dystrophin production. Lacking this protein, muscle fibers are sensitive to contraction-induced damage, resulting in chronic inflammation, impaired regeneration and replacement of muscle tissue by fibrotic and adipose tissues. Patients become wheelchair-dependent around the age of 12, require assisted ventilation around the age of 20 and often die prematurely in the 2nd to 4th decade of life due to cardiac or respiratory complications. Currently, many therapeutic approaches are in preclinical and clinical development, and aim to restore the missing protein, to improve regeneration and muscle mass and/or reduce inflammation and fibrosis. However, thus far translating findings from preclinical models to patients is challenging, where findings in mouse models are often not fully recapitulated in DMD patients.

As we consider the body of existing data on two key (DMD) mouse models – the classic C57BL/10ScSn-*Dmd*^{*mdx*}/J (B110/*mdx*) [1] and the promising new

D2.B10-*Dmd*^{*mdx*}/J (D2/*mdx*) model [2, 3] – a need exists to clarify the field’s present understanding of best practices in preclinical research using these two models, in particular when to use which model and which outcome measures to select, depending on the objectives. Charley’s Fund, a foundation dedicated to accelerating the development of life-saving treatments for DMD, engaged leading experts from both academia and industry to establish a collaborative effort to address this issue. The founding objectives set forth included: 1) help establish a grounding perspective on the present state of the data; 2) align on best practices for selecting the best model for various types of interventions and appropriate outcome measures to assess therapeutic effects covering muscle function and pathology; 3) ensure that natural history of the new D2/*mdx* model is collected without gaps; and 4) help establish more rigorous selection criteria designed to improve decisions on when to graduate a therapy to human clinical trials.

It was acknowledged that success in achieving these objectives would require significant, ongoing work from the parties involved. As a key grounding step, however, a 1.5-day key stakeholder meeting

entitled “Of Mice and Measures: A Project to Improve How We Advance DMD Therapies to the Clinic” was organized in Paris, France on 2–3 October 2017 to discuss the available science around this topic and prioritize action steps to move the field forward to improved therapeutic testing.

PRE-WORKSHOP PREPARATION AND DATA COLLECTION

Charley’s Fund engaged several leading experts to facilitate this project in 2016, creating a Scientific Organizing Committee responsible for directing the project, including this workshop, in early 2017. This committee included Charley’s Fund personnel, academic and industry individuals who directly perform or oversee preclinical testing in their own laboratories, and a statistician experienced in handling complex datasets. Over the course of nearly a year leading up to the workshop itself, the members of the committee convened frequently by phone and email.

The goals of this project required the collection and organization of functional, histological and biomarker data on the B110/*mdx*¹ and D2/*mdx* murine models from several different sources. D2/*mdx* mice that were backcrossed for 5 generations (3) and intercrossed for more than 9 generations were used for most of the studies. In order to allow consolidation and comparison of data from different sources, a data collection template was developed and sent to each contributor. This template described the format of the data required, standardized the units of collection as much as possible, and provided directions to ensure that each contributor provided all necessary information.

Over a period of 3 months, data templates were populated by 10 contributors out of the several groups contacted by Charley’s Fund that work with D2/*mdx*. These included seven academic groups, two private industry groups, and one non-profit organization. The data included over 230 different outcome parameters collected on 650 individual mice and yielding nearly 18,000 individual data points for comparison (see Table 1).

Working groups were organized to assess the outcomes according to a few specialized areas and develop findings to be presented for discussion at the workshop. These working groups included genet-

ics, natural history of disease, histological markers, biomarkers, behavior and skeletal muscle strength outcomes, respiratory assessments, cardiac assessments, and therapeutic responses. Multiple rounds of analysis and discussion were done within each working group prior to the October 2017 meeting.

Based on the goals set out at the beginning of the project, several different statistical analyses were planned. These included an observational description of the data to assess the present state of the information, comparisons between the classic B110/*mdx* and the new D2/*mdx* model to define outcomes most sensitive to the disease model, and comparisons between therapeutic interventions to define outcomes most useful in the evaluation of future therapies. This ambitious project benefitted from the overall coordination of the processes by dedicated people from Charley’s Fund, the organized way in which the data were collected, and having the statistical analysis for each working group performed by the same statistician.

Many statistical models were used in the analysis of these data and they are not discussed in detail here, however, a few overarching principles were followed. Descriptive statistics were presented collectively and stratified by mouse strain and by data source. For some outcomes collected on mice over a wide range of ages, descriptive statistics were also stratified by age. For inferential statistics, normality of all continuous outcomes was assessed before analysis and data transformations were applied where applicable. If the normality assumption could not be met, non-parametric statistical tests were used. All analyses of repeated assessments made on the same mouse accounted for the dependence of the assessments using mixed effects linear regression models.

Several general observations were apparent from the start.

1. Many outcomes were collected from only one source, making the assessment of inter-laboratory reliability not yet possible for those outcomes. Many laboratories focus on one particular area of assessment, thus the outcomes submitted by each source primarily involved one body system and these differed between sources.
2. Several outcomes were assessed by multiple sources but different methodologies were used for collection. This was particularly true with muscle strength assessments, where data collection was done *in vivo* and *ex vivo* by different sources.
3. For behavioral outcomes the variability was higher, indicating the need to consider

¹Of 650 individual mice, 29 *mdx* mice had a C57BL/6 background.

Table 1
Working groups and number of data points collected for this study

Class of readout	Readout	Data Data	Age range (weeks)	Strain measured				Contributors
				B110/ <i>mdx</i>	B110/ <i>wt</i>	D2/ <i>mdx</i>	D2/ <i>wt</i>	
Natural History	Lean and fat mass	394	6–12			*	*	Pfizer
	Body weight	2800	4–65	*	*	*	*	Aartsma-Rus, De Luca, Duan, Jax, Nagaraju, Pfizer, Sarepta, Spencer
	Tissue weights (excluding heart)	2135	4–65	*	*	*	*	Aartsma-Rus, De Luca, Duan, Nagaraju, Pfizer
	Tibia length	18	23–30			*	*	Duan
Functional	Wheel (voluntary and exhaustion)	318	8–58	*	*	*	*	De Luca, Nagaraju
	Open field (all assessments) ¹	376	8–65		*	*	*	Jax, Nagaraju, Spencer
	Grip strength (incl. normalized & hang test)	4142	4–65	*	*	*	*	Aartsma-Rus, De Luca, Jax, Nagaraju, Sarepta, Spencer
	Rotarod	612	6–65	*	*	*	*	Jax, Nagaraju, Sarepta
	Hindlimb sonography ²	72	28	*	*	*	*	De Luca
	Force (all muscles, all conditions)	1816	8–65	*	*	*	*	De Luca, Duan, Jax, Nagaraju, Pfizer
Cardiac	Functional assessments ³	761	16–58			*	*	De Luca, Duan, Nagaraju
	Size/weight assessments	288	4–65			*	*	Aartsma-Rus, Duan, Nagaraju
Histology	All assessments ⁴	1240	8–78			*	*	Aartsma-Rus, Duan, Jax, Nagaraju, Yokata
Respiratory	All assessments	1174	4–34			*	*	Aartsma-Rus, Spencer
Therapeutics	Gene expression in response to exon skipping	71	13	*		*		Aartsma-Rus
	CK	574	8–64	*	*	*	*	Aartsma-Rus, De Luca, Jax, Nagaraju
	LDH	76	9–64	*	*	*	*	De Luca
	Serum biomarkers ⁵	37	13–36			*	*	Duan, Nagaraju
	Dystrophin level	12	13	*		*		Aartsma-Rus
Biomarkers	Gene expression ⁶	939	10–34	*	*	*	*	Aartsma-Rus, Jax

¹Includes open field distance run, Digiscan measurements, and # of rearings. ²Includes volume and % vascularization. ³Shortening fraction, ejection fraction, stroke volume, and cardiac output also measured in B110 mice. ⁴Fiber area, fibrosis area, and % calcification also measured in B110 mice. ⁵Includes WBCs, macrophages, and neutrophils. ⁶Gene expression values from 12 genes in two different tissues (gastrocnemius and diaphragm).

operator-related issues and the need to follow standard operating procedures.

- The age of the mice from different sources varied considerably making direct comparisons between sources problematic.
- Data were not generated specifically for this project. Rather it was collected by each source from its own research studies. Thus, the majority of data collected were from mice included as part of drug studies but untreated with any active agent. This included both mice that had received a vehicle or sham treatment and mice that were completely untreated. The similarity of these two groups is unknown.

Given all of these observations, we generated descriptive statistics for all outcomes and performed

inferential statistical analysis where appropriate. The working groups generated presentations which included the results, notable findings and discussion points; these were presented to the group of key stakeholders at the workshop. After each working group presentation, a description of known issues, identification of additional questions, and conclusions based on the data presented were discussed in depth. The highlights of each working group are described below.

GENETICS OF THE D2/*mdx* MODEL

The D2/*mdx* carries the *mdx* mutation on the DBA/2J genetic background, which poses several advantages over the B110/*mdx* model. Several aspects

of the pathology match DMD patients' pathology, like the increased EBD-membrane permeability, fibrosis and TGF- β signaling, the impaired satellite cell activation and reduced regeneration [3]. Full genotyping of the D2/*mdx* background is still pending. However, it is already known that polymorphisms on the genes for latent TGF- β -binding protein 4 (*Ltbp4*) and annexin A6 (*Anxa6*) [4–6] respectively account for 40% and 7% of the phenotype in a mouse model for limb girdle muscular dystrophy type 2C (*Sgcg*) mice on a DBA/2J genetic background. The other 50% is unknown. The *Ltbp4* and *Anxa6* polymorphisms act in concert to modify muscle repair [7] and are thought to underlie the increased fibrosis and impaired regeneration in the D2/*mdx* model, respectively, while a variation in the *Abcc6* gene is responsible for the observed calcifications. Notably, the *Ltbp4* gene is also a modifier of human muscular dystrophy [8]. During breeding, genetic drift needs to be taken into account. The Jackson Laboratory advises refreshing in-house colonies every third generation.

While some natural history studies have been initiated with these mice, a lot of additional work is required to obtain a detailed picture of the natural history of disease. Furthermore, to determine the usefulness of the D2/*mdx* model for DMD therapeutic developments, the translatability of outcomes from mouse to human needs to be established.

It is important to be aware that wild type DBA/2J mice often have cardiac fibrosis and calcification, which poses challenges when using the D2/*mdx* model to study cardiac pathology, because it is not clear to which extent these are related to the lack of dystrophin or due to the DBA/2J background.

NATURAL HISTORY OF D2/*mdx* MICE

Life spans for animal models with dystrophin loss have been compared [9]. For DMD patients, life span has improved considerably in the last decade, due to improved pulmonary and cardiac care. For the D2/*mdx* life span, data up to 60 weeks of age are available and more is being collected.

Data on body weight of D2/*mdx* mice were collected from eight different laboratories. Body weights of D2/*mdx* mice were clearly lower than for B110/*mdx* and always lower than the wild type DBA/2J mice, in contrast with B110/*mdx* mice that tended to weigh more than their wild type strain (C57Bl/10Scsn) at all ages (Fig. 1). There was a difference of up to 10

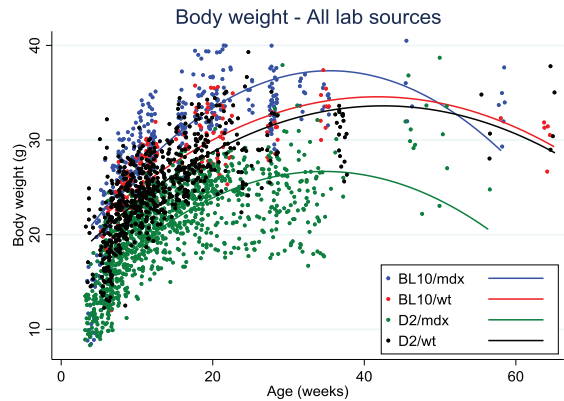


Fig. 1. Body weight distribution over age in B110/*mdx* (blue line), B110/wt (red line), D2/*mdx* (green line), D2/wt (black line). Data were collected from 8 labs.

grams between the body weights of D2/*mdx* from different sources that was seen with mixed genders as well as within males only. This variability may be due to husbandry, genetic drift or transportation stress if mice were purchased.

Skeletal muscle atrophy was observed in D2/*mdx* mice compared to DBA/2J wild types, which increased with age and was more apparent in bigger muscles than in small ones. D2/*mdx* mice had less fat and lean mass than DBA/2J wild types. Interestingly, spleen weight was higher in *mdx* as compared to wild type for both strains up to 30 weeks. Increased spleen weight probably reflects chronic muscle inflammation and muscle degeneration in these mice. Normalized heart weights were larger from a young age onwards in D2/*mdx* mice compared to DBA/2J wild type, and differences became more pronounced with increasing age. Functional heart deficits appeared earlier in D2/*mdx* than in B110/*mdx* and could possibly relate to the calcification process observed in all DBA/2J mice.

Serum creatine kinase (CK) levels decreased with age in both *mdx* mouse strains, similarly to what is observed in DMD patients [10]. This age-dependent decrease needs to be taken into account when CK is used to monitor the efficacy of a therapy that helps maintain muscle mass.

The forelimb grip strength normalized to body weight was higher in D2/*mdx* compared with B110/*mdx*. This may be due to the lower body weight observed in D2/*mdx* mice. The values of normalized grip strength tended to remain stable with age, and the difference to wild type was evident only for the B110 background. In DMD children, the loss of grip strength is more prominent than in *mdx* mice.

By contrast, the specific force, measured on the EDL muscle, was decreased in both *mdx* strains as compared to wild type, and B110/*mdx* specific force remains relatively constant with age although significantly lower than their wild type counterparts.

In summary, a certain variability was detected by comparing measurements from different labs, which may, next to husbandry conditions, be due to microbiota composition and nutritional aspects, instrument calibration and the genetic drift that occurs in the in-house colonies. In addition, the age range of D2/*mdx* mice likely to be influenced by treatments could not yet be defined.

HISTOLOGICAL MEASURES

Histology data were in general difficult to compare because labs collected different data and because the muscle pathology tended to be localized. The available data revealed that D2/*mdx* mice had fewer centronucleated myofibers than B16/*mdx* between 10–80 weeks of ages and that revertant fibers increased with age in B16/*mdx* but not in D2/*mdx* mice [11]. Both these parameters indicate less regeneration in D2/*mdx* compared with B16/*mdx* muscles and this was further supported by the D2/*mdx* mice having fewer eMHC positive fibers and BrdU positive nuclei. Terry Partridge presented data from his lab on BrdU incorporation as a measure of myogenic activity and showed that muscles D2/*mdx* began, like the original B110/*mdx*, to degenerate at 3 weeks of age, but did not start regenerating significantly over the subsequent two weeks, while B110/*mdx* had replaced 20–25% of muscle in the same period.

Fibrosis in D2/*mdx* was found to occur earlier in heart tissue (at 8 vs 36–40 weeks) and at higher levels in skeletal muscle tissue (for instance 20% vs 10% of total area at 10 weeks in the gastrocnemius) compared to B110/*mdx*. Calcification was more prominent in D2/*mdx* mice than in B110/*mdx* mice and both wild type strains. It affected skeletal muscles and heart to different extents and decreased in severity with age. The calcification was especially pronounced in the diaphragm of 8-week old D2/*mdx* mice, but reduced in area in adult (34 weeks) mice. This calcification at an early age (8 weeks) needs to be taken into account when measuring muscle function, but also when staining tissues, since it sequesters dyes and antibodies, resulting in very bright, aspecific staining of calcified areas. A longitudinal study and a better understanding of this pathological

event are needed. It was observed that calcification can be local and requires quantitation of the whole muscle cross-sectional area. Calcification does not occur in human dystrophic muscles, and over time fibrosis is replaced by adiposis in Duchenne muscles. In the first report describing D2/*mdx* mice, mice exhibit increased fat accumulation compared with B10/*mdx* mice [2]. However, no fat infiltration was observed in D2/*mdx* muscles in the frame of this work.

It is not known whether the reduced regeneration is due to an intrinsic problem of the satellite cells or to the elevated TGF β levels, or simply to decreased necrosis and increased fibrotic environment of the extracellular matrix. Increased levels of inflammation seen in D2/*mdx* [3] could also lead to impaired regeneration.

The use of Student's *t*-test to evaluate histological data and the tendency to consider histology as a quantitative measurement are questionable.

In general, it was noted that the picture of disease progression is not complete for D2/*mdx*: regeneration appears to be reduced in D2/*mdx* mice and fibrosis and calcification more prominent, especially at a young age (<12 weeks). At later time points the amount of fibrosis and calcification decreases, showing that regeneration does occur at some time. Clearly, further analysis is required to clarify the processes of necrosis, calcification, fibrosis and regeneration over time.

Miranda Grounds' group analysed the sciatic nerves innervating hind limb muscles, based on the hypothesis that the repeated cycles of myofibre necrosis and regeneration with associated disruption of neuromuscular junctions over many months, may result in denervation of myofibres with consequent changes in the sciatic nerve of dystrophic mice (it is noted that such changes become apparent in older normal mice after about 15 months of age (12)). Indeed, for male mice aged 13 months, the B110/*mdx* mice showed significant signs of denervation (tested as accumulation of S100 and tau protein in nerves), whereas this was neither seen in D2/*mdx* mice, nor in the wild type strains at this age. The simplest explanation is that the amount of ongoing necrosis of dystrophic muscles was much higher in the B110/*mdx*, compared with D2/*mdx*, mice. The new technology OPT (optical polarization tractography) allows to reconstruct organization of a whole fiber and may be useful to look deeper at all histological aspects, especially in mixed fast/slow muscles like the tibialis anterior [13, 14].

BIOMARKERS

D2/*mdx* mice showed higher inflammation than B110/*mdx* or wild type strains, as recorded by FACS analysis for total leucocyte and macrophage populations, inflammatory foci analysis and Cd68 qPCR, which decreased over time. The FACS analysis showed a CD3+ population in both D2/*mdx* and DBA/2J wild type at 10 weeks of age, but this population disappeared by 30 weeks of age. It was proposed that neutrophils should be quantified by peroxidase. Inflammation markers are relevant due to the current development of anti-inflammatory drugs, however, differences in mouse and human immune systems need to be considered, and the inflammatory process in D2/*mdx* is not yet fully understood.

Gene expression of fibrosis markers (*Tgf β* , connective tissue growth factor (*Ctgf*) and collagen) was increased in *mdx* compared to wild type and higher in D2/*mdx* than in B110/*mdx*, although not as high as one would expect based on histology. Levels were age dependent (values decreased in D2/*mdx* from 10 to 34 weeks) and muscle dependent, the diaphragm being more affected than the gastrocnemius. As mentioned, the time course of the fibrotic process, its interaction with calcification and the reason for the decline of TGF- β need to be better clarified in the D2/*mdx* mice.

Finally, CK values in blood were somewhat lower in D2/*mdx* than in B110/*mdx*, and tend to decrease with age, as also observed in DMD patients (Fig. 2). However, the variability is very high, and a range of sampling methods and protocols were used which precludes a direct comparison. It was suggested to test titin fragments in urine (mass-spectrometry, western blots, antibody stains) as an alternative marker for muscle pathology.

BEHAVIORAL ACTIVITIES AND SKELETAL MUSCLE FORCE MEASUREMENTS

The open field assessment records several parameters (total distance travelled, horizontal and vertical activity, number and time of movements, number of rearings and resting time) and showed consistent differences between *mdx* and wild type mice in both strains. However, the total distance travelled and the number of rearings was lower in D2/*mdx* than in B110/*mdx*. This strain-specific difference needs to be clarified and better assessed over time. Therefore, this

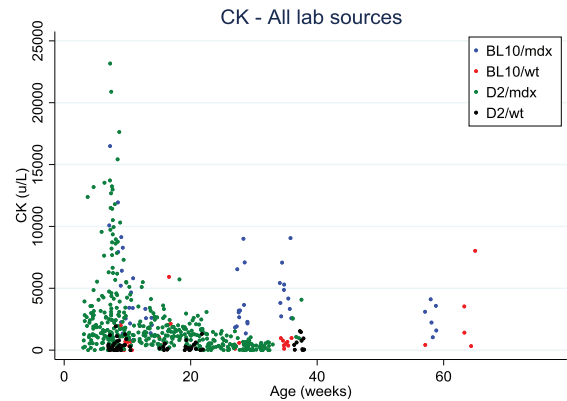


Fig. 2. Creatine kinase values distribution over age in B110/*mdx* (blue dots), B110/wt (red dots), D2/*mdx* (green dots), D2/wt (black dots). Data were collected from 4 labs.

assessment remains valuable for longitudinal, proof-of-concept studies but should be used in combination with behavioral, physiological and morphological measures in efficacy studies. The voluntary running wheel test and the exhaustion test with treadmill running showed differences in D2/*mdx* as compared to wild type, but there are still not enough data to support strain differences. The hanging wire test showed similar performances in *mdx* from both strains but a certain variability was present that could be partially explained by gender differences: male mice tend to perform less well than female mice and would be preferred in preclinical tests. Behavioral end points such as the hanging wire test are often influenced by many variables including sex, handling, time of testing etc.

As a general conclusion, fatigability tests are considered disease-relevant, given that in DMD patients, composite functional clinical readouts, such as 6-minute walk test or others, are accepted for approval of therapeutics. There are still a few issues that need further exploration, such as the progress of performance over time and possible correlations with reabsorption of calcifications and with fibrosis, the impact of exercise on pathology progression and the possible effects of drugs on the central nervous system with implications for volition and motivation. Furthermore, there is a clear need of consensus on best protocols to be used and on the appropriate method to acquire data: time-to-event representations on Kaplan-Meier diagrams may be more informative despite individual variability.

Grip strength declined with time for both *mdx* mouse strains as compared to wild type. However,

body weight increased in B110/*mdx* and decreased in D2/*mdx* mice over time making normalization to body weight a matter of debate, since it affects all normalizations of functional parameters. One option is to look at changes over time without normalization; another option, with a good sample size, is to adjust for the variant body weight and assess the functional parameter, if the variability is not too high. Therefore, although being a patient-relevant outcome, the grip strength is not recommended as primary outcome in mice, and should instead be used in combination with more direct muscle physiology assessments. Moreover, despite the high amount of data collected, the variability between labs and between operators was extremely high and will always represent an issue with this measurement. Grip strength remains an interesting parameter for longitudinal studies in mice, but ideally should be assessed by a single operator within an experiment.

Skeletal muscle force can also be measured directly with three methods: i) *In vitro/ex-vivo* method on isolated muscle, that assesses muscle performance without the interference of neural and vascular input; ii) *In situ* terminal method, that maintains physiological neural and vascular supply; iii) *In vivo* non terminal torque measurements, that record force from a group of muscles stimulated transdermally via the nerve and offers the advantages of maintaining physiological conditions, being minimally invasive and allowing multiple assessments in the same animal. For the *ex vivo* assays (i), *mdx* mice were weaker than wild type in both strains, and D2/*mdx* tended to be weaker than B110/*mdx* mice. EDL and diaphragm muscle force can be recommended as primary output, diaphragm data being useful for respiratory readouts. However, a parallel comparison of the two strains in the same laboratory over time is pivotal. For the *in situ* method (ii) on tibialis anterior, very limited data were available and no comparison between strains could be made. For the *in vivo* non-terminal torque assessment (iii), no standardized protocol is available; initial data suggest that D2/*mdx* were weaker than age-matched B110/*mdx*, but longitudinal studies will be needed. In general, muscle physiology assessments are considered valid as primary outcomes, they may detect drug effects that are not morphologically visible. However, further studies are needed to quantify the magnitude of strain difference, the age-dependent variations and the degree and progression of diaphragm impairment.

DIAPHRAGM AND RESPIRATORY PARAMETERS

Data pooling of respiratory function analyses was impossible due to the diversity in protocols used. In D2/*mdx* mice, the pulmonary function was reduced but the response to CO₂ challenge was not as strong as seen with B110/*mdx* mice, if the reduced weight of the D2/*mdx* strain was not accounted for. If the data were normalized, or if mice were tested at 32 weeks of age, this difference disappeared. This decline in function is relevant to DMD patients. However, the strain related differences need to be established. As described before, specific diaphragm force, measured by *ex vivo* analyses, decreased with age in B110/*mdx* and the drop in force was comparable in D2/*mdx* mice. This outcome is measured post-mortem and has a limited translatability, but is considered reproducible if standardized protocols are used.

The fibrosis and calcification of the diaphragm observed in D2/*mdx* mice were extensive and more pronounced than in other muscles: however, as in other muscles, the calcification decreased with age.

CARDIAC PARAMETERS

The most clinically relevant measurement on cardiac function is the shortening fraction, assessed by echocardiography in mice and patients; for mice, good standardized protocols (SOP: standardized operating procedure) exist and variability is small if these are followed. The ejection fraction is more difficult to measure clinically and in mice is a less reliable measure, most often calculated from M-mode measures [15, 16]. As measured using the TREAT-NMD/PPMD SOPs [17, 18], the shortening fraction was lower in D2/*mdx* than in its wild type counterpart aged 16–28 weeks, and this is normally seen in B110/*mdx* aged 36 to 52 weeks. However, the shortening fraction increased to normal levels at 52 weeks in the D2/*mdx* model [3]. The reason for this improvement in D2/*mdx* mice needs to be further elucidated, possibly heart wall thickness plays a role. Cardiac magnetic resonance (CMR) imaging was also used in the cardiac assessment of D2/*mdx* mice and demonstrated decreased function at 28 weeks of age [19]. Based on these measures, a mild cardiomyopathy occurs earlier in the D2/*mdx* compared to B110/*mdx* mice.

Other cardiac functional measures are: pressure/volume loops, (cardiac catheter assay) and ECG. The heart rate is increased in both *mdx* strains as compared to their wild type controls and is seen in patients with DMD. Fibrosis was evident in hearts of D2/*mdx* mice, as early at 8 weeks of age and, as expected, calcification was high in the 10-week-old heart of both *mdx* and wild type on DBA background, but absent on the B110 background. The etiology and significance of cardiac calcifications require further study. In summary, histological and biochemical assessments of fibrosis are recommended, but best methods need to be chosen. If possible, fibrosis quantification should be aligned with biomarkers or CMR of cardiac function or measurements of myocardial composition [19].

THERAPEUTIC INTERVENTIONS

In general, the D2/*mdx* mouse offers either a larger or smaller observational window to test effects of therapeutic interventions depending on the type of readout chosen. Therefore, which model is most relevant will differ based on the intervention that is studied.

Thus far, limited data are available on therapeutic interventions tested in the D2/*mdx* mouse model [20, 21]. A small pilot was performed comparing treatment with antisense oligonucleotides in D2/*mdx* and B110/*mdx* mice. There was no difference between B110/*mdx* and D2/*mdx* mice in exon skipping efficiency, suggesting that this model is suitable for future studies.

Natural history data are incomplete, the pathophysiology at early ages and the course of the processes of necrosis, regeneration, calcification and fibrosis need to be elucidated, which for now hampers designing optimal therapeutic intervention studies in the D2/*mdx* model. Such preclinical therapeutic studies should be run in parallel with vehicle-treated controls, and with a choice of outcomes not affected by intrinsic or operator-derived variability. Ideally, studies should be conducted in two independent labs. Treatments tested should target different pathways (for instance inflammation, muscle mass growth, dystrophin restoration, etc.).

RECOMMENDATIONS

From our efforts some points can be summarized:
1) Most reliable outcomes are muscle specific force

measures and open field; outcomes with caveats are grip strength and mouse fatigability, fibrosis and whole-body plethysmography with CO₂ challenge, cardiac outcomes; caution is needed with assays such as CK in blood and histological central nucleation. 2) Experimental conditions, especially mouse gender, power size per outcome, ages, vehicle use, best protocol but also husbandry and timing of measurements should be defined. 3) Blinded assessment is recommended especially when testing therapeutics in mice of the same genotype (undistinguishable by eyes) and especially for those methodologies which require a subjective role of the experimenter (i.e., behavioural tests, hand-made histology analysis, etc.). Additionally, data groups can be blinded for the statistical analysis. Existing TREAT-NMD standardized operating procedures can be selected and extended to the D2/*mdx* mouse model, others may need to be drafted ex novo.

It is recognized that a robust natural history data collection should represent a key achievement prior to testing therapies, and an appropriate location for a sharable database should be identified. In particular, more detailed longitudinal studies on the courses of necrosis/regeneration, calcification and fibrosis over time are needed; time points for histological data collection should be agreed upon (suggested: 2 wk, 4 wk, 6 wk, 12 wk, 6 mo, 1 year); cardiac pathology needs to be followed up after 24 weeks. For biomarkers, it seems more practical to choose clinically validated markers and test their preclinical translatability following treatment in mice. Furthermore, a minimum of criteria that need to be met for proof-of-concept versus preclinical studies should be defined.

As next steps, the following priorities were identified for the field and will be considered by the Organizing Committee. First, the existing TREAT-NMD SOPs need to be reviewed to assess their suitability for D2/*mdx* analysis and, if suitable, D2/*mdx* data need to be integrated. SOPs currently missing should be identified, and working groups to generate them should be established. Second, for the outcome measures identified as usable and reliable, data gaps should be filled. Ideally, the data should be collected in a concerted and independent experimental effort of laboratories using standardized and complementary approaches. Interested stakeholders will be needed for funding to support this time- and resource-intensive action. Such an effort will help to enable the compilation of a natural history database that hosts datasets for D2/*mdx* and B110/*mdx* mice to

benefit the scientific community (both academic and private) interested in translational research in DMD. Finally, considering the fact that even these expert groups collected data for outcomes in different ways, it is clear that the field needs to reach consensus on which outcomes to select and how to assess them. A white paper is under discussion to identify and define the best practices.

ACKNOWLEDGMENTS

A special thanks to the Scientific Organizing Committee: Annemieke Aartsma-Rus, Annamaria De Luca, Heather Gordish-Dressman, Kanneboyina Nagaraju, Melissa Spencer, Maaïke van Putten, and Laura Dalle Pазze and Arati Kreibich of Charley's Fund for the coordination of all efforts before and during the workshop.

The workshop was sponsored by Charley's Fund in collaboration with the TREAT-NMD Alliance. Dr. Nagaraju is supported by National Institutes of Health grants # K26OD011171. Dongsheng Duan is supported by NIH (AR67985, AR69085, NS90634), DOD (MD150133), Jesse's Journey, Jackson Freed DMD Research Fund and Solid Biosciences.

DISCLOSURES

Dongsheng Duan is a member of the scientific advisory board for Solid Biosciences and an equity holder of Solid Biosciences. The Duan lab has received research support from Solid Biosciences. Annamaria De Luca has provided ad hoc consultancy in 2017 to Akashi Therapeutics and Grünenthal GmbH. Frederick J. Schnell and Marco A. Passini are paid employees of Sarepta Therapeutics Inc. Laurent Bogdanik and Cathleen Lutz are paid employees of The Jackson Laboratory. David Elsey is a paid employee of Summit Therapeutics plc. Mahasweta Girgenrath is a paid employee of Pfizer Inc. J. Patrick Gonzalez is a paid employee of Solid Biosciences. Andy Nichols is a paid employee of Catabasis Pharmaceuticals Inc. Zhong Zhong is a paid employee of Wave Life Sciences Ltd. Toshifumi Yokota has provided ad hoc consultancy in 2018 to Agada Biosciences.

REFERENCES

- [1] Bulfield G, Siller WG, Wight PA, Moore KJ. X chromosome-linked muscular dystrophy (mdx) in the mouse. *Proc Natl Acad Sci U S A*. 1984;81(4):1189-92.
- [2] Fukada S, Morikawa D, Yamamoto Y, Yoshida T, Sumie N, Yamaguchi M, et al. Genetic background affects properties of satellite cells and mdx phenotypes. *Am J Pathol*. 2010;176(5):2414-24.
- [3] Coley WD, Bogdanik L, Vila MC, Yu Q, Van Der Meulen JH, Rayavarapu S, et al. Effect of genetic background on the dystrophic phenotype in mdx mice. *Hum Mol Genet*. 2016;25(1):130-45.
- [4] Heydemann A, Ceco E, Lim JE, Hadhazy M, Ryder P, Moran JL, et al. Latent TGF-beta-binding protein 4 modifies muscular dystrophy in mice. *J Clin Invest*. 2009;119(12):3703-12.
- [5] Swaggart KA, Demonbreun AR, Vo AH, Swanson KE, Kim EY, Fahrenbach JP, et al. Annexin A6 modifies muscular dystrophy by mediating sarcolemmal repair. *Proc Natl Acad Sci U S A*. 2014;111(16):6004-9.
- [6] Swaggart KA, McNally EM. Modifiers of heart and muscle function: Where genetics meets physiology. *Exp Physiol*. 2014;99(4):621-6.
- [7] Quattrocelli M, Capote J, Ohiri JC, Warner JL, Vo AH, Earley JU, et al. Genetic modifiers of muscular dystrophy act on sarcolemmal resealing and recovery from injury. *PLoS Genet*. 2017;13(10):e1007070.
- [8] Flanigan KM, Ceco E, Lamar KM, Kaminoh Y, Dunn DM, Mendell JR, et al. LTBP4 genotype predicts age of ambulatory loss in Duchenne muscular dystrophy. *Ann Neurol*. 2013;73(4):481-8.
- [9] McGreevy JW, Hakim CH, McIntosh MA, Duan D. Animal models of Duchenne muscular dystrophy: From basic mechanisms to gene therapy. *Dis Model Mech*. 2015;8(3):195-213.
- [10] Hathout Y, Brody E, Clemens PR, Cripe L, DeLisle RK, Furlong P, et al. Large-scale serum protein biomarker discovery in Duchenne muscular dystrophy. *Proc Natl Acad Sci U S A*. 2015;112(23):7153-8.
- [11] Rodrigues M, Echigoya Y, Maruyama R, Lim KR, Fukada SI, Yokota T. Impaired regenerative capacity and lower revertant fibre expansion in dystrophin-deficient mdx muscles on DBA/2 background. *Sci Rep*. 2016;6:38371.
- [12] Krishnan VS, White Z, McMahon CD, Hodgetts SI, Fitzgerald M, Shavlakadze T, et al. A Neurogenic Perspective of Sarcopenia: Time Course Study of Sciatic Nerves From Aging Mice. *J Neuropathol Exp Neurol*. 2016;75(5):464-78.
- [13] Wang Y, Zhang K, Wasala NB, Duan D, Yao G. Optical polarization tractography revealed significant fiber disarray in skeletal muscles of a mouse model for Duchenne muscular dystrophy. *Biomed Opt Express*. 2015;6(2):347-52.
- [14] Wang Y, Zhang K, Duan D, Yao G. Heart structural remodeling in a mouse model of Duchenne cardiomyopathy revealed using optical polarization tractography [Invited]. *Biomed Opt Express*. 2017;8(3):1271-6.
- [15] Spurney CF, McCaffrey FM, Cnaan A, Morgenroth LP, Ghelani SJ, Gordish-Dressman H, et al. Feasibility and Reproducibility of Echocardiographic Measures in Children with Muscular Dystrophies. *J Am Soc Echocardiogr*. 2015;28(8):999-1008.
- [16] Soslow JH, Xu M, Slaughter JC, Stanley M, Crum K, Markham LW, et al. Evaluation of Echocardiographic Measures of Left Ventricular Function in Patients with Duchenne Muscular Dystrophy: Assessment of Reproducibility and Comparison to Cardiac Magnetic Resonance Imaging. *J Am Soc Echocardiogr*. 2016;29(10):983-91.
- [17] Available from: <http://www.treat-nmd.eu/research/preclinical/dmd-sops>.

- [18] Duan D, Rafael-Fortney JA, Blain A, Kass DA, McNally EM, Metzger JM, et al. Standard Operating Procedures (SOPs) for Evaluating the Heart in Preclinical Studies of Duchenne Muscular Dystrophy. *J Cardiovasc Transl Res.* 2016;9(1):85-6.
- [19] Vohra R, Batra A, Forbes SC, Vandenborne K, Walter GA. Magnetic Resonance Monitoring of Disease Progression in mdx Mice on Different Genetic Backgrounds. *Am J Pathol.* 2017;187(9):2060-70.
- [20] Ito T, Ogawa R, Uezumi A, Ohtani T, Watanabe Y, Tsujikawa K, et al. Imatinib attenuates severe mouse dystrophy and inhibits proliferation and fibrosis-marker expression in muscle mesenchymal progenitors. *Neuromuscul Disord.* 2013;23(4):349-56.
- [21] Takemoto Y, Inaba S, Zhang L, Baba K, Hagihara K, Fukada S. An herbal medicine, Go-sha-jinki-gan (GJG), increases muscle weight in severe muscle dystrophy model mice. *Clinical Nutrition Experimental.* 2017;16:13-23.