Early insights from 'Of Mice and Measures', a collaborative project to improve models and methods for preclinical research in Duchenne muscular dystrophy, and its first focus on the D2.B10-*Dmd*^{mdx}/J (D2/*mdx*) and C57BL/10ScSn-*Dmd*^{mdx}/J (Bl10/*mdx)* mouse models

Laura Dalle Pazze[1], Heather Gordish-Dressman[2]

[1]Charley's Fund, New York, NY, USA and [2]Children's National Medical Center, Washington, DC, USA

Abstract

Optimizing preclinical tools and methods to evaluate therapeutic candidates is critical to improve decision-making about advancing therapies to clinical testing. In collaboration with leading experts in the neuromuscular community and TREAT-NMD network, Charley's Fund, a patient-founded research nonprofit, organized an effort to address a timely topic of this nature: how to utilize a promising newer mouse model, the D2.B10-*Dmd*^{mdx}/J (D2/*mdx*), which has been speculated to recapitulate human pathology better than the commonly used C57BL/10ScSn-*Dmd*^{mdx}/J (Bl10/*mdx*). The group convened working groups and collected and analyzed D2/*mdx* and Bl10/*mdx* data from multiple academic and industry sources. Nearly 18,000 data points on mutant and WT strains were gathered spanning 10 labs worldwide, 650 individual mice, and 230 different functional, histological, imaging, biochemical and molecular parameters. In addition to information about best practices and data gaps in the two models, the cross-lab data comparison yielded findings with broader implications. Even in labs of leading experts, notable differences in practices and results were identified. To begin, no single outcome was assessed across all 10 labs. Most commonly assessed was body weight in 6 labs, followed by serum creatine kinase in 5 labs. Eight other measures were assessed in 3 labs, and another 16 in 2 labs. The effort also revealed: a) labs often use different protocols for the same measure; b) behavioral assessments in particular yield variable outcomes; and c) inconsistencies exist in control arm design and use of vehicle, sham, and untreated mice — for example 1/4 control mice in the aggregate dataset received vehicle but 3/4 received no treatment. This project highlights an important need to understand and address these differences to improve consistency, quality, and coordination of preclinical research in DMD — and a compelling opportunity to intervene early to organize effective research using the new D2/*mdx* model.

Background

- Charley's Fund was founded in 2004 to accelerate drug development for DMD
- The current DMD research landscape features an increasingly active pipeline of therapies; accordingly, particular importance exists around developing research tools and improved processes that benefit all development efforts
- Research-focused nonprofits like Charley's Fund —with sole focus on driving benefit to patients — are particularly poised to address opportunities of this nature
- A particular challenge in DMD is translatability of preclinical animal data to human clinical trials: best practices and better methods are urgently needed
- Charley's Fund engaged experts from academia and industry to develop a program to:
  - a) help establish a grounding perspective on the present state of the data
  - b) align on best practices for selecting the best model and measures
  - c) ensure data on natural history of the new D2/*mdx* model is complete
  - d) establish rigorous criteria to graduate a therapy to human clinical trials

Methods

Ongoing since fall 2016, 'Of Mice and Measures' is a collaborative initiative among academic, industry, and nonprofit partners in the DMD research community. Charley's Fund serves as a central coordinating party and works collaboratively with a Scientific Organizing Committee and contributing partners to identify opportunities, develop strategies, convene contributors, and undertake action steps. An initial October 2017 workshop in Paris provided a key grounding step. It is acknowledged that success in achieving the overall initiative's objectives will require significant, ongoing work from the parties involved.

10 labs across the world contributed data to the initial workshop. It is hoped that more will add their data as the effort continues.

Key steps around the initial 2017 workshop

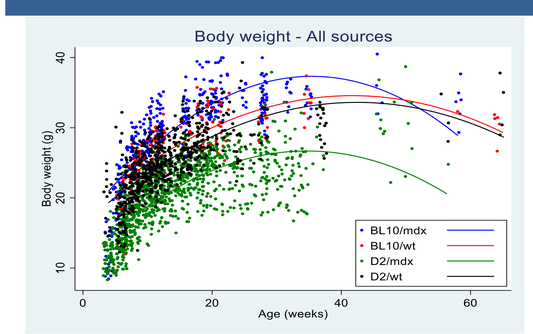
- Fall 2016: Charley's Fund and founding experts begin discussion and initial due diligence
- Winter 2017: Scientific Organizing Committee convened
- Spring 2017: Data template designed, call for contributions issued, data collected
- Summer 2017: Working groups formed, multiple rounds of analysis, discussion, and insight generation conducted
- October 2017: Initial 1.5 day workshop convened
- Winter-Spring 2018: Results documented, workshop report generated, next steps outlined and initiated

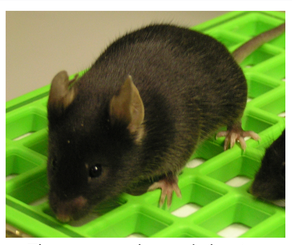
Data Collection

Outcome class	Outcome	Data Points	Age range (weeks)	Strain Measured: BL10/mdx	BL10/wt	DB2/mdx	DB2/wt	Contributors
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 (including 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, Yokota
Respiratory	All assessments	1174	4 - 34			*	*	Aartsma-Rus, Spencer
Therapeutics	Gene expression in response to Exon skipping	71	13	*		*		Aartsma-Rus
Biomarkers	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
	Gene expression[6]	939	10 - 34	*	*	*	*	Aartsma-Rus, Jax

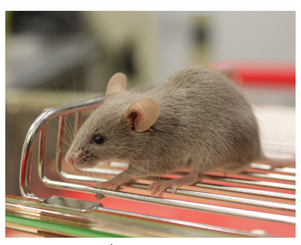
[1]Includes open field distance run, Digiscan measurements, and # of rearings. [2]Includes volume and % vascularization. [3]Shortening fraction, ejection fraction, stroke volume, and cardiac output also measured in Bl10 mice. [4]Fiber area, fibrosis area, and % calcification also measured in Bl10 mice. [5]Includes WBCs, macrophages, and neutrophils. [6]Gene expression values from 12 genes in two different tissues (gastrocnemius and diaphragm).

Natural History





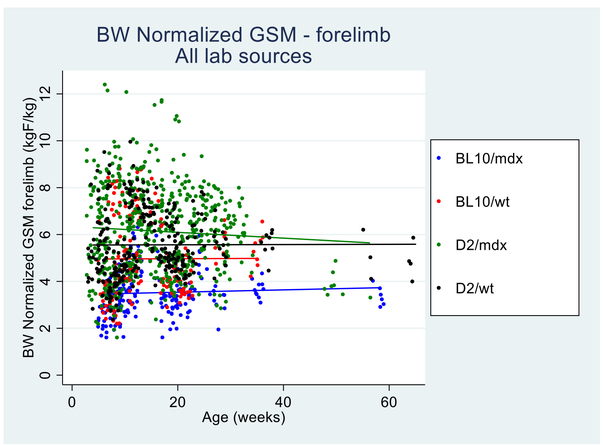
The commonly used classic C57BL/10ScSn-*Dmd*^{mdx}/J (Bl10/*mdx*) mouse model



The newer D2.B10-*Dmd*^{mdx}/J (D2/*mdx*) mouse model

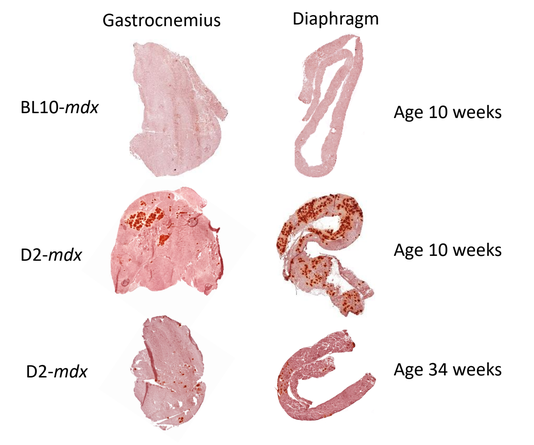
This figure clearly shows a difference in body weight between Bl10 and D2 mice. However it also shows a difference of up to 10g in D2/*mdx* mice from different sources. This variability could be attributed to several factors including the husbandry at each facility, genetic drift in the mice, or stress from transportation. If we observe up to a 10g difference between mice of the same strain, detecting a difference due to treatment effects becomes very challenging.

Behavioral Measures



Here we see a difference in body-weight normalized grip strength among mice of different strains. We also see a very wide range within the same strain. Some of this variability can be attributed to the body weight differences shown; however the assessment of grip strength itself shows a high degree of variability. This variability can have several sources including the element of subjectivity in measurement and the choice of assessment (i.e., maximum value or average value of several repetitions).

Histological Measures

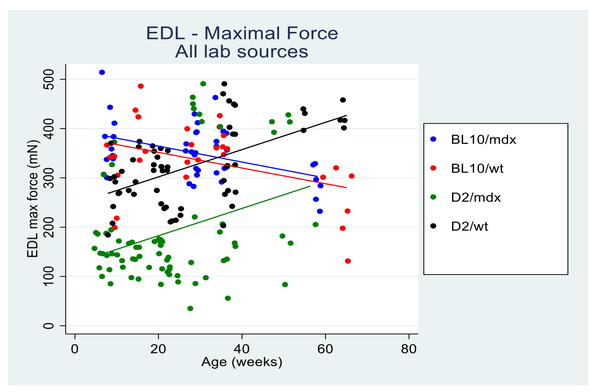


Histological assessments such as inflammation, fibrosis, degeneration, regeneration, and central nucleation are not consistently measured in all the labs. One of the striking histological features of the D2/*mdx* model is extensive calcifications in skeletal muscle. Here we show an assessment of calcifications in skeletal muscle and the diaphragm that are prominent early (10 weeks) but spontaneously resolve as the mice age (34 weeks). Calcification is not a major histological feature of Bl10/*mdx* mice.

Recommendations

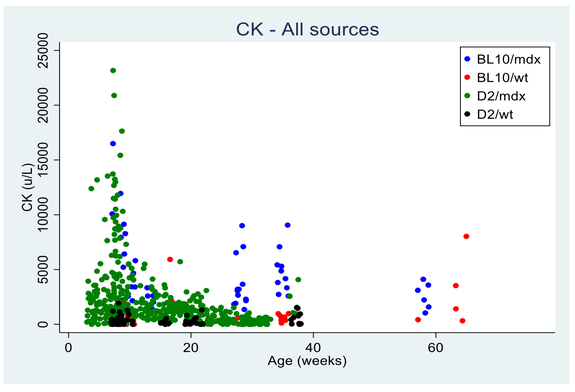
- Investigators should follow established SOPs for the collection of data to decrease variability due to collection method
- Investigators should choose outcomes based on ability to assess efficacy given the mouse model and mechanism of action of the treatment, rather than the ease or comfort with collection method
- Investigators should follow blinding procedures, especially for outcomes that have a component of subjectivity involved
- The statistical analysis of preclinical data should be performed thoughtfully. This includes:
  - Using nonparametric tests where warranted
  - Using an analysis method that is appropriate for the outcome's distribution
  - When assessing an outcome over time or age, use methods that account for repeated assessments and evaluate non-linear relationships
  - Clearly define the dependent and independent variable(s)
  - Use time-to-event analysis methods where appropriate

Functional Measures



This figure shows overall differences between mouse strains, a decrease in maximal force with age in Bl10 mice, yet an increase with age in D2 mice. The variability, however, makes distinguishing between strains at a given age difficult. At 38 weeks of age, force values from all four strains completely overlap. Any treatment expected to show a strain difference at 38 weeks would be problematic and would require a very large sample size.

Biomarkers



Here we see measured values of CK in each strain with a very high degree of variability with values ranging from near 0 to 25000 u/L in young mice. While CK levels are known to have a wide range, different procedures for collecting serum at different facilities likely contribute to the extremes we see here. This variability makes CK a challenging outcome. In addition, the inherent variability in CK levels typically requires nonparametric statistics for analysis.

Conclusions

The first step in Of Mice and Measures provided insights on:

(1) Findings and recommendations that can be implemented immediately to improve how the DMD research community utilizes the D2/*mdx* and Bl10/*mdx* models to evaluate therapies
(2) Broader findings about opportunities to gather missing data, to identify best practices to be standardized, and to ensure effective communications to disseminate important information

Initial findings are available in detail in a workshop report in *Journal of Neuromuscular Diseases*. Additional priority next steps identified and underway include:

- Documentation of SOPs for D2/*mdx* mice*;* updates to existing SOPs for Bl10/*mdx* mice
- A follow-up, more data-heavy publication with new data and additional analyses
- A natural history study to fill in identified gaps across D2/*mdx* and Bl10/*mdx* mice
- Improved summary guidance materials for DMD mice and preclinical protocols
- Consideration of future efforts such as a single searchable DMD mouse database and broader minimal preclinical data standards in DMD