

New semi-automated tool for the quantitation of MR imaging to estimate *in vivo* muscle disease severity in mice

Emily A. Waters¹, Chad R. Haney¹, Lauren. A Vaught², Elizabeth M. McNally², and Alexis R. Demonbreun^{2,3*}

¹Chemistry of Life Processes Institute and Biomedical Engineering, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA

²Center for Genetic Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA

³Department of Pharmacology, Northwestern University Feinberg School of Medicine, Chicago, IL 60611, USA

*To whom correspondence should be addressed:

Alexis R. Demonbreun PhD

Center for Genetic Medicine

Northwestern University

303 E Superior SQ 5-512

Chicago, IL 60611 USA

T: 1 312 503 6249 F: 1 312 503 5603

alexis.demonbreun@northwestern.edu

Abbreviations: Duchenne Muscular Dystrophy (DMD), magnetic resonance imaging (MRI)

Keywords: Duchenne Muscular Dystrophy, MRI, muscle, mdx

Abstract

The pathology in Duchenne muscular dystrophy (DMD) is characterized by degenerating muscle fibers, inflammation, fibro-fatty infiltrate, and edema, and these pathological processes replace normal healthy muscle tissue. The *mdx* mouse model is one of the most commonly used preclinical models to study DMD. Mounting evidence has emerged illustrating that muscle disease progression varies considerably in *mdx* mice, with inter-animal differences as well as intra-muscular differences in pathology in individual *mdx* mice. This variation is important to consider when conducting assessments of drug efficacy and in longitudinal studies. Magnetic resonance imaging (MRI) is a non-invasive method that can be used qualitatively or quantitatively to measure muscle disease progression in the clinic and in preclinical models. Although MR imaging is highly sensitive, image acquisition and analysis can be time intensive. The purpose of this study was to develop a semi-automated muscle segmentation and quantitation pipeline that can quickly and accurately estimate muscle disease severity in mice. Herein, we show that the newly developed segmentation tool accurately divides muscle. We show that measures of skew and interdecile range based on segmentation sufficiently estimate muscle disease severity in healthy wildtype and diseased *mdx* mice. Moreover, the semi-automated pipeline reduced analysis time by nearly 10-fold. Use of this rapid, non-invasive, semi-automated MR imaging and analysis pipeline has the potential to transform preclinical studies, allowing for pre-screening of dystrophic mice prior to study enrollment to ensure more uniform muscle disease pathology across treatment groups, improving study outcomes.

Introduction

Duchenne muscular dystrophy (DMD) is an inherited muscle wasting disease caused by loss of dystrophin. Dystrophin, a large protein that links the cytoskeleton with the extracellular matrix, is encoded by the *DMD* gene. In the absence of dystrophin, the muscle membrane becomes easily disrupted which causes muscles to degenerate, and in its place, the intact muscle becomes replaced by fibrosis and fat, as well as immune infiltrate. The dystrophin-deficient *mdx* mouse model is routinely used in preclinical therapeutic trials as it recapitulates many of the same pathological features seen in humans with DMD. The *mdx* mouse model on the C57Bl/10 background harbors a naturally occurring mutation in exon 23 of dystrophin resulting in a premature stop codon ^{1,2}. Although the *mdx* mouse does not produce any detectable dystrophin protein, this model exhibits only mild histological and functional deficits in comparison to humans with DMD. A more severe mouse model referred to as *mdxD2* was generated through backcrossing the *mdxC57Bl/10* model over 5 generations onto the DBA/2J strain ³. The muscle pathology in *mdxD2* mice has more similar features to human DMD pathology with increased areas of muscle damage, fibrofatty deposition, impaired repair capacity, and, along with these findings, greater functional deficits compared to the *mdxC57Bl/10* model ⁴⁻⁸. The increase in disease progression on the DBA/2J background has been linked to multiple genetic modifiers, including Latent TGFB binding protein 4 (*Ltpb4*), annexin A6 (*Anxa6*), and osteopontin (*Spp1*) ⁹⁻¹¹.

The availability of multiple mouse models of DMD has allowed for extensive preclinical studies evaluating potential therapeutics for the treatment of DMD. With preclinical testing occurring across multiple animal colonies, there have been efforts to standardize procedures to increase the reliability and translatability of endpoint measures (TREAT-NMD.org protocols). It is well known that the *mdx* model displays a wide range of both intra-animal and inter-animal variability ¹². For example, *mdx* littermates can exhibit virtually no muscle damage or fibrosis in one subject, while another is severely affected. Moreover, within the same animal one muscle may have extensive pathology while the contralateral muscle is minimally affected. This inherent variation can contribute to poor assay sensitivity and necessitate large cohorts of animals to achieve the statistical power required to show treatment effects ¹³. Spurney et al showed that quantitative histology measures of degeneration and regeneration and creatine kinase showed high variance in the *mdx* model. Evans' blue dye uptake, as a measure of muscle membrane leak, also showed both animal to animal variability and intra-animal variability across studies ¹⁴.

Understanding and mitigating factors that can influence preclinical outcome measures is critical when designing and evaluating therapeutic efficacy studies. Environmental factors such as cage design, light/dark cycle, food, sex, age, and weight can be more easily controlled than biological factors. Non-invasive imaging techniques such as magnetic resonance imaging (MRI) can provide information on *in vivo* muscle tissue health without the need for tissue sampling or sacrifice. Although MRI can accurately distinguish between healthy and diseased tissue, qualitative estimates of severity are highly subjective, and quantitative image analysis can be cumbersome, making it less than ideal as a screening tool. Further, it can be challenging to identify a numerical summary statistic that reflects qualitative differences readily observed between imaging datasets. For example, in datasets with small, localized regions of signal change, simple signal averaging can obscure differences due to the much larger areas of unchanged tissue. Indeed, previous efforts to use average T_2 values as an imaging biomarker of disease progression in *mdx* mice have been hampered by high variance that has been interpreted as a lack of significant difference between timepoints or groups, but more likely reflects the inadequacy of bulk averaging as a summary statistic¹⁵. We propose a rapid high-throughput screening method that uses a combination of MR imaging, semi-automated segmentation, and quantitative analysis with meaningful summary statistics to estimate disease

severity *in vivo*, balance treatment groups, and *a priori* exclude animals with unusually high or low disease burdens.

Methods

Animals

Male mice, 6-7 weeks old, WT, *mdxB10* and *mdxD2* mice were purchased from the Jackson Laboratory (stock 000664, 001801, and 013141). Mice were housed in a specific pathogen-free facility on a 12-hour light/12-hour dark cycle and fed ad libitum in accordance with the Northwestern University's Institutional Animal Care and Use Committee regulations.

MR Image acquisition

Mice were anesthetized in an induction chamber with inhaled isoflurane and then transferred to a dedicated imaging bed with a nosecone to deliver continuous isoflurane. Each mouse was positioned prone with legs tucked beneath the abdomen to reduce susceptibility artifacts. Respiration was monitored using a pressure sensitive pillow and warming was performed. MRI was performed on a 9.4T Bruker Biospec 9430 (Bruker Corporation, Billerica, MA, USA) with a 30 cm bore and 12 cm gradient insert, running Paravision 6.0.1. Each mouse was in a 40 mm quadrature volume coil (Bruker) operating in transmit/receive mode. After localizer images were acquired, a T_2 map was acquired using a spin echo sequence (Multi Slice Multi Echo, MSME) oriented axially and centered at the mid-calf. The following parameters were used: TR = 4000 ms, TE = 9-225 ms (30 echoes, echo spacing = 9 ms), MTX = 256 x 256, FOV 3.5 x 3.5 cm, 5 slices, 1 mm slice thickness and 1 signal average. Acquisition time was approximately 18 minutes.¹⁶

Training of Segmentation Network

Images were imported into Amira 2020.2 software (Thermo Fisher Scientific, Waltham, MA, USA) and the first echo of the T_2 map acquisition (TE = 9 ms), a relatively proton-density weighted image, was used to segment a region of interest (ROI) containing hindlimb and paraspinal muscles, and remove other tissues such as bladder, skin, fat, and chemical shift artifacts. As this was a time-consuming manual process, a deep learning prediction model was trained in Amira using the built-in tools. A training dataset was assembled using segmentation and image data 24 MDX scans and 7 wildtype scans and concatenated with a validation dataset assembled from segmentation and image data obtained by rescanning 5 MDX mice and 3 wildtype mice 2-3 days after their initial imaging session. A preliminary model was trained using the "Deep Learning Training" module in Amira with type BackbonedUNet, number of classes = 2, resnet18 backbone, patch size 128, batch size 8, max patch overlap 0.3, maximum number of epochs 100, Adam optimizer, learning rate 0.0001, and no data augmentation. A refined model was trained using the same module, initialized with the weights of the preliminary model, with the following modifications: patch size 256, maximum number of epochs 500, learning rate 0.0005, and data augmentation with geometry transforms including horizontal and vertical flip, 30-degree rotation, 10% zoom, and 10 degree shear, and an early stopping criterion to stop if no model improvement for 25 iterations. The "Segmentation Metrics" Amira Xtra module was used to assess the performance of the model on the validation dataset.¹⁷

Semiautomated Image Segmentation

After the deep learning model was generated, segmentation was performed by importing the first echo of the T_2 map acquisition into Amira and generating a segmentation using the "Deep Learning Prediction" module using the established model weights. The segmentation was inspected by a trained observer and manually refined as necessary to correct minor errors (e.g.,

incomplete removal of the bladder). Muscle ROIs were exported from Amira as a mask image to be used in further processing steps.

MR Image Processing

All echoes of the T_2 map acquisition, were imported into JIM 7 (Xinapse Systems Ltd, West Bergholt, Essex, UK), masked to include only muscle voxels using the mask image generated in the segmentation step, and fit using the built in nonlinear curve fitting module. A custom Python script was used to extract a list of muscle T_2 values from the resultant maps. T_2 values < 15 ms and > 100 ms were excluded as these overwhelmingly corresponded to voxels with low signal and low-quality fitting.

Evans Blue Dye

To validate the skew and interdecile range of the T_2 distribution as metrics for disease severity, they were compared to Evans Blue dye uptake. WT mice (n=5), *mdx* mice (n=4) and *mdxD2* mice (n=3) underwent the MRI protocol described above and then underwent measurements of Evans Blue dye uptake that were quantified as described previously^{9, 18, 19}. Briefly, mice were injected with 5 μ l per gram of body weight of 10 μ M Evans Blue dye (E2129, Sigma-Aldrich). Mice were euthanized, tissue excised, minced, and placed in 1ml of formamide in a 24-well plate placed at 55 °C for 2 hrs. Subsequently, absorbance was measured at 620 nm on a Synergy HTX multi-mode plate reader (BioTek®) from 200 μ l of sample solution in a 96-well plate. Each sample was assessed in duplicate. Results are reported as the average arbitrary optical density units/mg of tissue.

Statistical Analysis

A key purpose of this study was to identify and validate relevant summary statistics for MRI image distributions based on the hypothesis that the distribution of T_2 values would be sufficiently non-normal as to render mean and standard deviation inappropriate summary statistics. A custom Python script was used to test the normality of the distribution of T_2 values for each animal using D'Agostino's K2 test, and compute summary statistics.²⁰⁻²⁴ All distributions observed were markedly non-normal ($p < 0.001$). As a result, summary statistics of skew and interdecile range were calculated to describe the width and asymmetry of distributions and enable quantitative comparisons. Correlation coefficients for the test-retest variability and the comparisons of skew and IDR with Evans' Blue dye uptake were calculated using a simple linear regression model.

Code Availability

Imaging data and associated code for processing and analysis is available online.²⁵

Results

Development of a semi-automated muscle segmentation tool. Manual segmentation of the muscle ROI including the hindlimb and paraspinal muscles was a lengthy manual process due to the need to manually remove bone signals as well as signals from normal fluid and fat that are present subcutaneously and at intramuscular interfaces. The signals from regions of fat additionally required removal of chemical shift artifacts. Because of the manual and subjective nature of this process, it was highly subject to intra- and interoperator variability. Use of the deep learning model reduced the time for generating muscle ROIs from over an hour to a few seconds and reduced total hands-on analyst time to 5-10 minutes. In the validation dataset, the model-predicted label field matched the ground truth label field with a Jaccard Index of 0.97,

indicating excellent agreement. An example segmentation of a dataset not included in the training data is shown in **Figure 1**.

Stratification of muscle disease through semi-automated MR imaging analysis. *mdxB10* mice expressing a severe disease phenotype (**Figure 2A**), mild disease phenotype (**Figure 2B**), and healthy wildtype controls (**Figure 2C**) were imaged to cover a wide range of muscle health. Examination of the T_2 maps shown in **Figure 2A** and **Figure 2B** illustrates the wide range of disease phenotype present in this cohort of same-aged and same-sex *mdx* mice. **Figure 2A** illustrates a mouse with severe disease as indicated by extensive regions of mildly elevated T_2 , in addition to several focal regions of highly elevated T_2 . In contrast, **Figure 2B** illustrates a mouse with comparatively mild disease, in which there are focal regions with mildly elevated T_2 , and many fewer voxels with highly elevated T_2 . The healthy wildtype mouse shown in **Figure 2C** has few voxels with high T_2 and these are primarily located at the intermuscular interfaces. For each T_2 map, the distribution of T_2 values is plotted immediately to the right, with the mean plotted in red and the standard deviation represented by red error bars (**Figure 2A-C**). The distribution of T_2 values is immediately seen to be non-normal (as confirmed by D'Agostino's K^2 test of normality). A plot of skew vs. interdecile range of the distribution of T_2 values was effective in stratifying disease severity and correlated well with visual assessment of the selected T_2 maps (**Figure 2D**). Together these data show that MR image-based measures extracted using the newly developed, semi-automated segmentation and analysis pipeline stratify muscle disease severity across healthy to severe pathology.

Semi-automated MR image-based measures are reproducible across disease severity. To assess the repeatability of measurements of the skew and interdecile range of the T_2 distribution, 5 *mdx* mice and 3 WT mice were reimaged two days after their initial imaging session. Both sets of data were analyzed by a trained observer according to the protocol described above. Both skew and interdecile range were found to be highly repeatable across imaging sessions (**Figure 3**). The first and second measurements were highly correlated for both skew ($R^2 = 0.97$, $p < 0.001$) and interdecile range (IDR) ($R^2 = 0.81$, $p = 0.002$)

High correlation between measures of muscle membrane leak and semi-automated MR imaging-based measures. Quantification of Evans blue dye uptake is commonly used to evaluate muscle membrane permeability and stability as a measure of muscle health, as dye is excluded from healthy intact muscle. To determine if the MR image-based measures, skew, and IDR, correlate with dye uptake measures, *mdxB10*, *mdxD2*, and WT mice underwent MR imaging and subsequently were evaluated for Evans blue dye uptake in hindlimb muscles. Good agreement was observed between the *in vivo* imaging based metrics (skew and IDR) and the *ex vivo* Evans Blue dye uptake measurements (**Figure 4**). Healthy WT muscle had the lowest level of dye uptake and IDR followed by *mdxB10*, with *mdxD2* muscle taking up the highest levels of dye and having the highest IDR values. IDR exhibited a strong linear correlation with the Evans Blue Dye with $R^2 = 0.9$ ($p < 0.001$). Skew was also linearly correlated with $R^2 = 0.71$ ($p < 0.001$). This suggests that the *in vivo* imaging based metrics can serve as a surrogate measure of muscle membrane health and integrity as compared to the terminal Evans blue dye uptake assessment.

Discussion

We have outlined a non-invasive, quick, and reproducible imaging and analysis protocol that can accurately stratify animals to aid in preclinical therapeutic study design. Screening of animals to assess disease severity prior to assigning treatment groups can help mitigate the potential effects of the wide range of disease severities in the *mdx* model, as treatment effects can be obscured if there are differences in the underlying severity of disease between the groups. Many studies rely on comparing histopathological findings or Evans Blue dye uptake in separate groups of untreated and treated animals. These measures rely on terminal procedures that cannot be undertaken at the outset of a study.

Previous efforts to use MRI in the assessment of disease severity in *mdx* mice have focused on T_2 maps and T_2 weighted images because regions of muscle damage tend to have abnormally elevated T_2 values. However, researchers have consistently noted very small differences in the average value of muscle T_2 (on the order of a few milliseconds) between *mdx* and wildtype animals.^{15, 26} Such a small difference, while in some cases statistically significant, can be easily confounded by minor systemic errors in image acquisition and fitting. Upon deeper interrogation of our study data, we found that the mean and standard deviation of T_2 values over a region of interest comprising hindlimb and paraspinal musculature were very poor descriptors of the actual distribution of measured T_2 values and were not appropriate for stratifying the severity of disease in *mdx* animals. Attempts to differentiate the groups by setting a T_2 threshold and counting voxels above that threshold were highly dependent on the specific threshold selected and ineffective at stratifying disease.

Inspection of the distribution of T_2 values for a muscle region of interest in *mdx* and WT mice yielded stark differences in the peak shape and width of the distribution of T_2 values between the *mdx* and WT mice. In particular, in WT mice, the distribution of T_2 values was a sharp narrow peak roughly centered around a low T_2 value, with a shallow tail of higher values corresponding to blood vessels, inter-muscular interfaces, and noisy voxels with poor T_2 fitting. In contrast, for the *mdx* mice, the distribution was shifted rightwards in the direction of higher T_2 values, and markedly asymmetric with a “heavy” tail of high T_2 values, mostly corresponding to regions of edema and necrosis. The distributions were assessed for normality and found to be markedly non-normal, confirming that the mean and standard deviation were inappropriate summary statistics for these data.

Because the goal was to identify animals with a higher concentration of high- T_2 voxels, corresponding to the heaviness of the distribution tail, skew was investigated as a higher-order summary statistic and interdecile range as a measure of spread of the values, on the premise that these metrics would better reflect the shape of the distribution and might be more suitable for stratifying degree of disease. The plot of these two metrics correlated with the spectrum of disease severity, clearly and repeatably differentiating *mdx* animals from wildtype animals and correlating well with the terminal measurements made using Evans Blue Dye. In this study, skew and IDR were identified as metrics suitable to describe the acquired data. In future studies, this analysis will be extended to a more comprehensive radiomics approach in an effort to identify texture-based image features that may reflect localized areas of muscle damage.²⁷

A major strength of this rapid semiautomated imaging and analysis pipeline is that images of several major hindlimb muscle groups can be acquired in approximately 30 minutes, with only 10-15 minutes per animal needed for semiautomated segmentation and processing. This significantly increases the throughput relative to manual segmentation and processing, which required 1.5-2 hours per animal, and renders it suitable for screening of cohorts of animals prior to assigning treatment groups. The proposed pipeline could substantially improve the quality of studies of therapeutic efficacy in mouse models of muscular dystrophy by ensuring that study cohorts are appropriately balanced across the spectrum of disease and a priori excluding animals with either extremely severe or extremely mild disease.

AUTHOR CONTRIBUTIONS

LV and AD performed the mouse harvest and EBD study. EW and CH performed the MR imaging, segmentation, and analysis. EW, CH, AD and EM conceived and designed the studies, analyzed the data, and wrote the manuscript.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health AR052646, HL167813, NS047726, NS127383, Additional funding was through Lakeside Discovery. MRI was performed at the Northwestern University Center for Advanced Molecular Imaging (RRID:SCR_021192) generously supported by NCI CCSG P30 CA060553 awarded to the Robert H Lurie Comprehensive Cancer Center. EAW is supported by grant number 2020-225578 from the Chan Zuckerberg Initiative DAF, an advised fund of Silicon Valley Community Foundation.

CONFLICT OF INTEREST

EMM has consulted for Amgen, AstraZeneca, Cytokinetics, Pfizer, Tenaya Therapeutics and is the founder of Ikaika Therapeutics. ARD is the Chief Scientific Officer at Ikaika Therapeutics. These activities are unrelated to the content of this manuscript. The other authors have no conflicts of interest to declare.

Figures and Legends

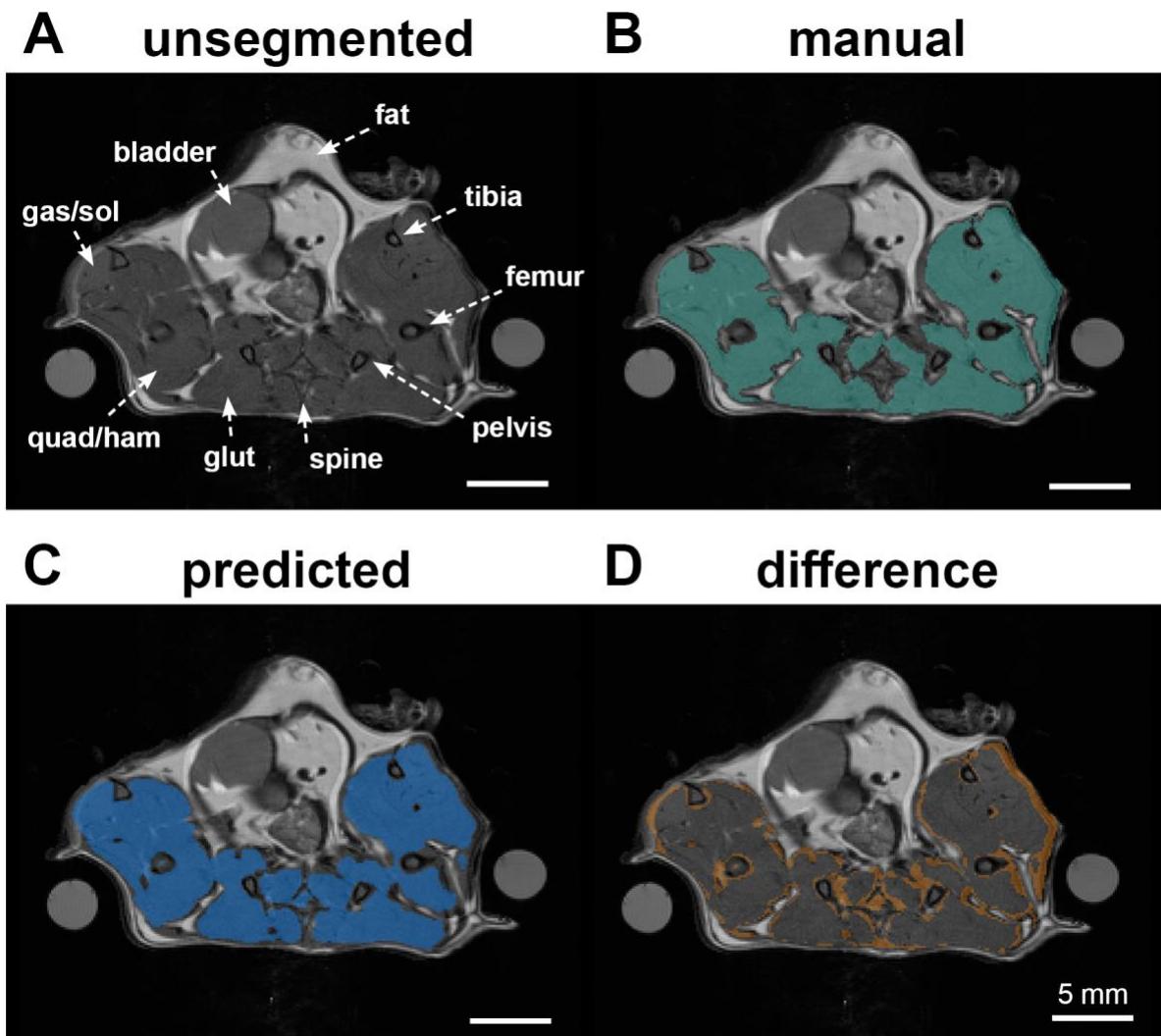


Figure 1: Reliable semi-automatic muscle segmentation. (A) Proton density weighted axial image showing hindlimb and paraspinal musculature. Muscle groups visualized include gastrocnemius and soleus (lower hindlimb), quadriceps and hamstrings (upper hindlimb) and gluteal muscles (near pelvis). (B) Manual segmentation of muscle ROI. (C) Result of segmentation using the trained deep learning model. (D) Difference between manual and automated segmentation showing overall excellent agreement with minor variation around the muscle-skin and muscle-bone interfaces.

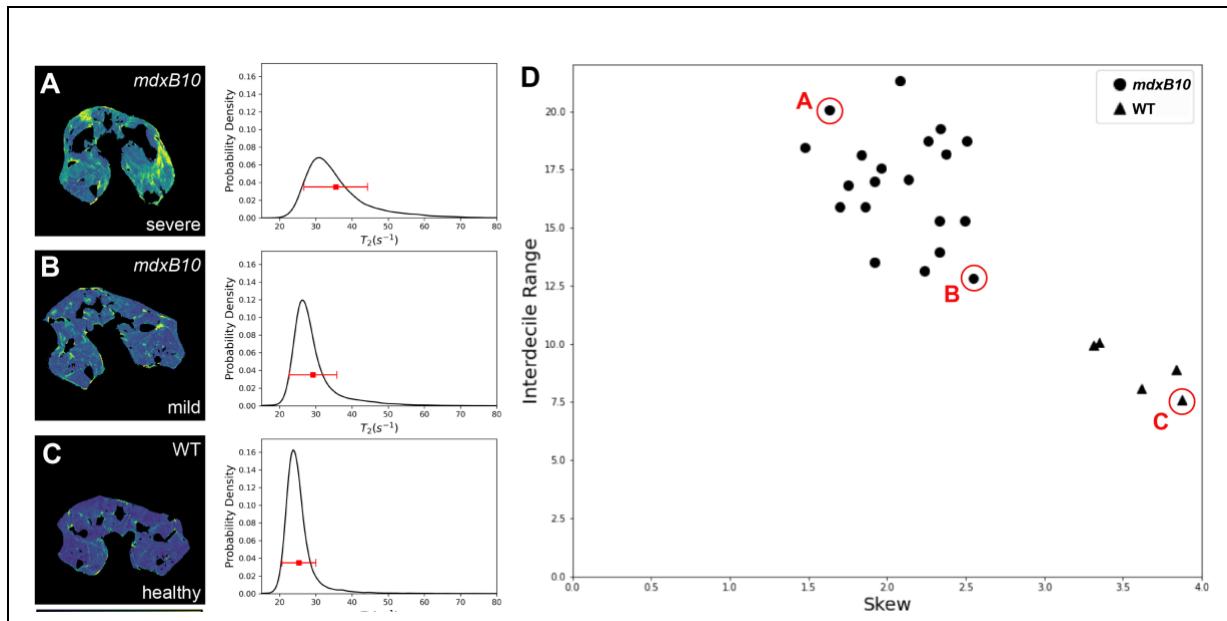


Figure 2: MR image-based measures stratify muscle disease severity. (A-C) on the left are representative T_2 map slices from 3 different animals in the study. (A) shows an *mdxB10* animal with a severe muscle disease phenotype as evidenced by the globally higher T_2 values as well as localized areas of high T_2 corresponding to muscle damage. (B) shows an *mdxB10* animal with a relatively mild disease phenotype seen by moderate regions of high T_2 values. (C) is a healthy wildtype (WT) control animal showing minimal regions of high T_2 aside from normal intramuscular interfaces. Plots to the right of each image show the distribution of T_2 values for that animal. The red square and error bars correspond to the mean and standard deviation of the T_2 values and demonstrate the insensitivity of those metrics to changes in the shape of the distribution. (D) A plot of skew vs interdecile range (IDR) for 20 *mdxB10* animals and 5 healthy WT control animals, demonstrating that skew and IDR clearly separate the two genotypes as well as stratifying disease severity for the *mdx* animals. Data points marked A-C correspond to the individual datasets shown on the left.

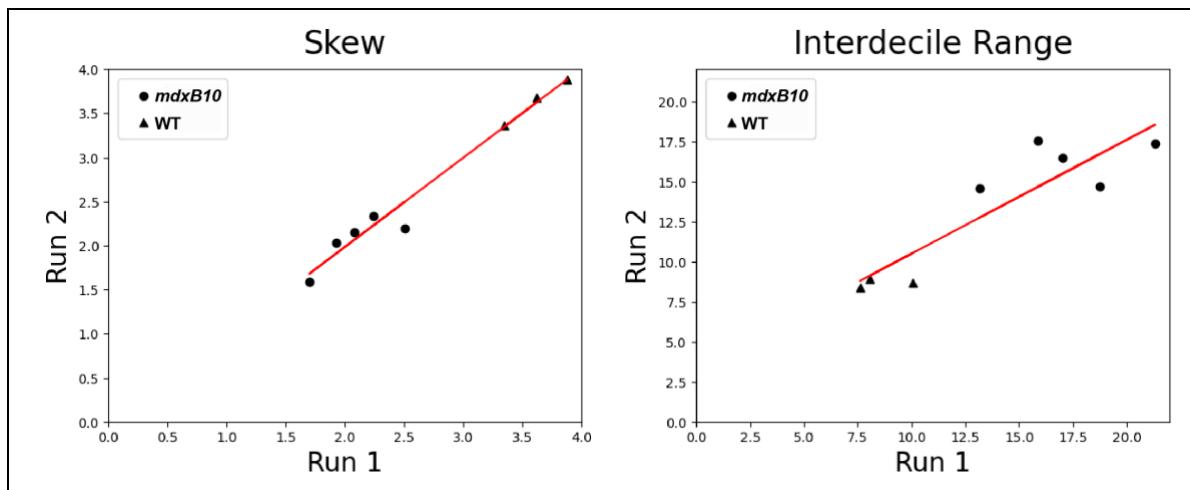


Figure 3: Reproducible MR image-based measures in healthy and diseased muscle. Test-retest reliability comparison of measured values of skew (left) and IDR (right) in *mdxB10* mice reimaged 2 days after the initial imaging session. Both values exhibited a strong linear correlation with skew $R^2 = 0.97$ and IDR $R^2 = 0.81$ including diseased (mdx) and healthy (WT) muscle.

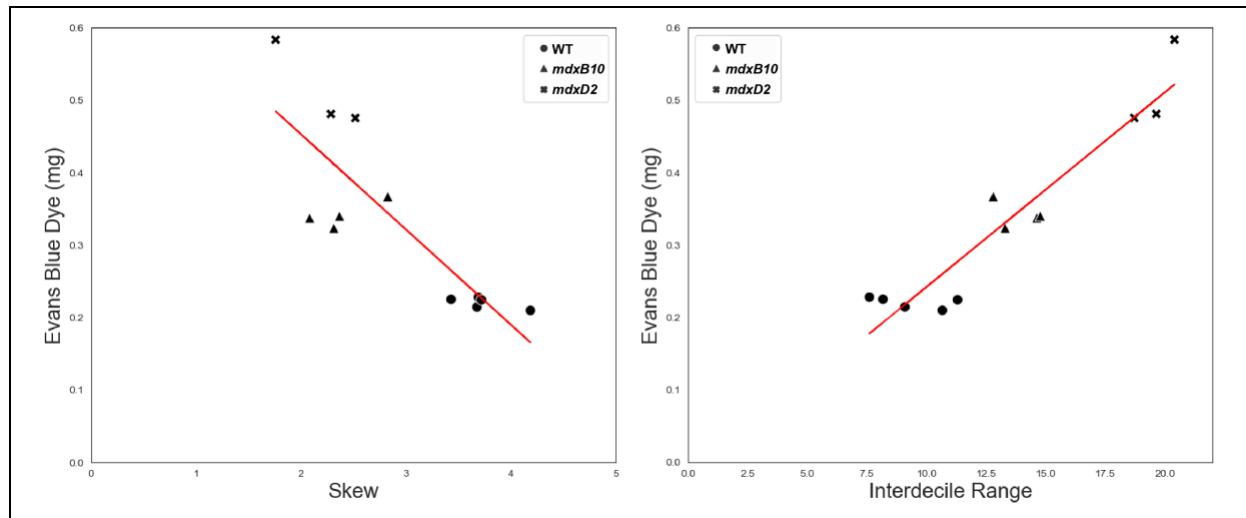


Figure 4: Strong correlation between dye uptake and MRI measures across muscle disease severity. Comparison of noninvasive *in vivo* image-based measurements of disease severity (skew and interdecile range) with ex vivo measurement of Evans Blue Dye. No muscle disease occurred in healthy WT controls (n=5), mild disease was present in *mdxB10* mice (n=4), and severe disease was evident in *mdxD2* mice (n=3). Both measures exhibited a strong linear correlation with the Evans Blue Dye (R^2 vs skew: 0.71; R^2 vs IDR: 0.9).

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. doi: 10.1073/pnas.81.4.1189. PubMed PMID: 6583703; PMCID: PMC344791.
2. Sicinski P, Geng Y, Ryder-Cook AS, Barnard EA, Darlison MG, Barnard PJ. The molecular basis of muscular dystrophy in the mdx mouse: a point mutation. *Science.* 1989;244(4912):1578-80. doi: 10.1126/science.2662404. PubMed PMID: 2662404.
3. Fukada S, Morikawa D, Yamamoto Y, Yoshida T, Sumie N, Yamaguchi M, Ito T, Miyagoe-Suzuki Y, Takeda S, Tsujikawa K, Yamamoto H. Genetic background affects properties of satellite cells and mdx phenotypes. *Am J Pathol.* 2010;176(5):2414-24. Epub 20100319. doi: 10.2353/ajpath.2010.090887. PubMed PMID: 20304955; PMCID: PMC2861106.
4. Capote J, Kramerova I, Martinez L, Vetrone S, Barton ER, Sweeney HL, Miceli MC, Spencer MJ. Osteopontin ablation ameliorates muscular dystrophy by shifting macrophages to a pro-regenerative phenotype. *J Cell Biol.* 2016;213(2):275-88. Epub 20160418. doi: 10.1083/jcb.201510086. PubMed PMID: 27091452; PMCID: PMC5084275.
5. Hammers DW, Hart CC, Matheny MK, Wright LA, Armellini M, Barton ER, Sweeney HL. The D2.mdx mouse as a preclinical model of the skeletal muscle pathology associated with Duchenne muscular dystrophy. *Sci Rep.* 2020;10(1):14070. Epub 20200821. doi: 10.1038/s41598-020-70987-y. PubMed PMID: 32826942; PMCID: PMC7442653.
6. Mazala DA, Novak JS, Hogarth MW, Nearing M, Adusumalli P, Tully CB, Habib NF, Gordish-Dressman H, Chen YW, Jaiswal JK, Partridge TA. TGF-beta-driven muscle degeneration and failed regeneration underlie disease onset in a DMD mouse model. *JCI Insight.* 2020;5(6). Epub 20200326. doi: 10.1172/jci.insight.135703. PubMed PMID: 32213706; PMCID: PMC7213798.
7. van Putten M, Putker K, Overzier M, Adamzek WA, Pasteuning-Vuhman S, Plomp JJ, Aartsma-Rus A. Natural disease history of the D2-mdx mouse model for Duchenne muscular dystrophy. *FASEB J.* 2019;33(7):8110-24. Epub 20190401. doi: 10.1096/fj.201802488R. PubMed PMID: 30933664; PMCID: PMC6593893.
8. Coley WD, Bogdanik L, Vila MC, Yu Q, Van Der Meulen JH, Rayavarapu S, Novak JS, Nearing M, Quinn JL, Saunders A, Dolan C, Andrews W, Lammert C, Austin A, Partridge TA, Cox GA, Lutz C, Nagaraju K. Effect of genetic background on the dystrophic phenotype in mdx mice. *Hum Mol Genet.* 2016;25(1):130-45. Epub 20151112. doi: 10.1093/hmg/ddv460. PubMed PMID: 26566673; PMCID: PMC4690497.
9. Heydemann A, Ceco E, Lim JE, Hadhazy M, Ryder P, Moran JL, Beier DR, Palmer AA, McNally EM. Latent TGF-beta-binding protein 4 modifies muscular dystrophy in mice. *J Clin Invest.* 2009;119(12):3703-12. Epub 20091102. doi: 10.1172/JCI39845. PubMed PMID: 19884661; PMCID: PMC2786802.
10. Quattrocelli M, Capote J, Ohiri JC, Warner JL, Vo AH, Earley JU, Hadhazy M, Demonbreun AR, Spencer MJ, McNally EM. Genetic modifiers of muscular dystrophy act on sarcolemmal resealing and recovery from injury. *PLoS Genet.* 2017;13(10):e1007070. Epub 20171024. doi: 10.1371/journal.pgen.1007070. PubMed PMID: 29065150; PMCID: PMC5669489.
11. Swaggart KA, Demonbreun AR, Vo AH, Swanson KE, Kim EY, Fahrenbach JP, Holley-Cuthrell J, Eskin A, Chen Z, Squire K, Heydemann A, Palmer AA, Nelson SF, McNally EM. Annexin A6 modifies muscular dystrophy by mediating sarcolemmal repair. *Proc Natl Acad Sci U S A.* 2014;111(16):6004-9. Epub 20140409. doi: 10.1073/pnas.1324242111. PubMed PMID: 24717843; PMCID: PMC4000833.
12. Grounds MD, Radley HG, Lynch GS, Nagaraju K, De Luca A. Towards developing standard operating procedures for pre-clinical testing in the mdx mouse model of Duchenne muscular dystrophy. *Neurobiol Dis.* 2008;31(1):1-19. Epub 20080409. doi: 10.1016/j.nbd.2008.03.008. PubMed PMID: 18499465; PMCID: PMC2518169.

13. Spurney CF, Gordish-Dressman H, Guerron AD, Sali A, Pandey GS, Rawat R, Van Der Meulen JH, Cha HJ, Pistilli EE, Partridge TA, Hoffman EP, Nagaraju K. Preclinical drug trials in the mdx mouse: assessment of reliable and sensitive outcome measures. *Muscle Nerve*. 2009;39(5):591-602. doi: 10.1002/mus.21211. PubMed PMID: 19260102; PMCID: PMC4116326.
14. Kobayashi YM, Rader EP, Crawford RW, Campbell KP. Endpoint measures in the mdx mouse relevant for muscular dystrophy pre-clinical studies. *Neuromuscul Disord*. 2012;22(1):34-42. Epub 20111210. doi: 10.1016/j.nmd.2011.08.001. PubMed PMID: 22154712; PMCID: PMC3264796.
15. 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. doi: 10.1016/j.ajpath.2017.05.010. PubMed PMID: 28826559; PMCID: PMC5809503.
16. Waters EA. MR imaging of the mouse hindlimb musculature (T2 weighted and T2 map): *protocols.io*; 2023. Available from: <https://protocols.io/view/mr-imaging-of-the-mouse-hindlimb-musculature-t2-we-cucjwsun>.
17. Image Segmentation Evaluation Using Standard Metrics (Python): Thermo Fisher Scientific; 2021 [updated March 8, 2021; cited 2023 May 1, 2023]. Available from: <https://xtras.amira-avizo.com/xtras/image-segmentation-evaluation-using-standard-metrics-python>.
18. Demonbreun AR, Wyatt EJ, Fallon KS, Oosterbaan CC, Page PG, Hadhazy M, Quattrocelli M, Barefield DY, McNally EM. A gene-edited mouse model of limb-girdle muscular dystrophy 2C for testing exon skipping. *Dis Model Mech*. 2019;13(2). Epub 20191104. doi: 10.1242/dmm.040832. PubMed PMID: 31582396; PMCID: PMC6906631.
19. Demonbreun AR, Fallon KS, Oosterbaan CC, Bogdanovic E, Warner JL, Sell JJ, Page PG, Quattrocelli M, Barefield DY, McNally EM. Recombinant annexin A6 promotes membrane repair and protects against muscle injury. *J Clin Invest*. 2019;129(11):4657-70. doi: 10.1172/JCI128840. PubMed PMID: 31545299; PMCID: PMC6819108.
20. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, Burovski E, Peterson P, Weckesser W, Bright J, van der Walt SJ, Brett M, Wilson J, Millman KJ, Mayorov N, Nelson ARJ, Jones E, Kern R, Larson E, Carey CJ, Polat İ, Feng Y, Moore EW, VanderPlas J, Laxalde D, Perktold J, Cimrman R, Henriksen I, Quintero EA, Harris CR, Archibald AM, Ribeiro AH, Pedregosa F, van Mulbregt P, Vijaykumar A, Bardelli AP, Rothberg A, Hilboll A, Kloeckner A, Scopatz A, Lee A, Rokem A, Woods CN, Fulton C, Masson C, Häggström C, Fitzgerald C, Nicholson DA, Hagen DR, Pasechnik DV, Olivetti E, Martin E, Wieser E, Silva F, Lenders F, Wilhelm F, Young G, Price GA, Ingold G-L, Allen GE, Lee GR, Audren H, Probst I, Dietrich JP, Silterra J, Webber JT, Slavič J, Nothman J, Buchner J, Kulick J, Schönberger JL, de Miranda Cardoso JV, Reimer J, Harrington J, Rodríguez JLC, Nunez-Iglesias J, Kuczynski J, Tritz K, Thoma M, Newville M, Kümmeler M, Bolingbroke M, Tartre M, Pak M, Smith NJ, Nowaczyk N, Shebanov N, Pavlyk O, Brodtkorb PA, Lee P, McGibbon RT, Feldbauer R, Lewis S, Tygier S, Sievert S, Vigna S, Peterson S, More S, Pudlik T, Oshima T, Pingel TJ, Robitaille TP, Spura T, Jones TR, Cera T, Leslie T, Zito T, Krauss T, Upadhyay U, Halchenko YO, Vázquez-Baeza Y, SciPy C. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature Methods*. 2020;17(3):261-72. doi: 10.1038/s41592-019-0686-2.
21. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*. 2011;12(85):2825-30.
22. McKinney W. Data Structures for Statistical Computing in Python. In: van der Walt S, Millman J, editors. *Proceedings of the 9th Python in Science Conference*; 20102010. p. 56-61.

23. Hunter JD. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*. 2007;9(3):90-5. doi: 10.1109/MCSE.2007.55.
24. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, Wieser E, Taylor J, Berg S, Smith NJ, Kern R, Picus M, Hoyer S, van Kerkwijk MH, Brett M, Haldane A, del Río JF, Wiebe M, Peterson P, Gérard-Marchant P, Sheppard K, Reddy T, Weckesser W, Abbasi H, Gohlke C, Oliphant TE. Array programming with NumPy. *Nature*. 2020;585(7825):357-62. doi: 10.1038/s41586-020-2649-2.
25. Waters EA. Waters Demonbreun MDX MRI Manuscript Data Package. 1.0.0 ed. Prism: Galter Health Sciences Library, Northwestern University; 2023.
26. Dunn JF, Zaim-Wadghiri Y. Quantitative magnetic resonance imaging of the mdx mouse model of Duchenne muscular dystrophy. *Muscle Nerve*. 1999;22(10):1367-71. doi: 10.1002/(sici)1097-4598(199910)22:10<1367::aid-muss5>3.0.co;2-h. PubMed PMID: 10487902.
27. Apte AP, Iyer A, Crispin-Ortuzar M, Pandya R, van Dijk LV, Spezi E, Thor M, Um H, Veeraraghavan H, Oh JH, Shukla-Dave A, Deasy JO. Technical Note: Extension of CERR for computational radiomics: A comprehensive MATLAB platform for reproducible radiomics research. *Med Phys*. 2018. Epub 20180613. doi: 10.1002/mp.13046. PubMed PMID: 29896896; PMCID: PMC6597320.