



HOKKAIDO UNIVERSITY

Title	A Systematic Histopathological Study of Duchenne Muscular Dystrophy Using Semi-Quantitative Image Analysis, Digital Restoration Techniques, and Exploratory Statistical Approaches
Author(s)	山角, 哲弘
Degree Grantor	北海道大学
Degree Name	博士(医学)
Dissertation Number	甲第17035号
Issue Date	2026-06-30
DOI	https://doi.org/10.14943/doctoral.k17035
Doc URL	https://hdl.handle.net/2115/100088
Type	doctoral thesis
File Information	YAMAKADO_Tetsuhiro.pdf, 全文



学 位 論 文

A Systematic Histopathological Study of Duchenne Muscular Dystrophy Using Semi-Quantitative Image Analysis, Digital Restoration Techniques, and Exploratory Statistical Approaches

(デュシェンヌ型筋ジストロフィー筋生検標本の探索的統合解析による組織病理学研究)

2026 年 6 月

北 海 道 大 学

山 角 哲 弘

学 位 論 文

A Systematic Histopathological Study of Duchenne Muscular Dystrophy Using Semi-Quantitative Image Analysis, Digital Restoration Techniques, and Exploratory Statistical Approaches

(デュシェンヌ型筋ジストロフィー筋生検標本の探索的統合解析による組織病理学研究)

2026 年 6 月

北 海 道 大 学

山 角 哲 弘

Index

List of Publications and Conference Presentations-----	1
Abstract-----	2
List of Abbreviations -----	6
1. Introduction -----	10
2. Materials and Methods -----	12
2.1 Patients -----	12
2.1.1 Standard Protocol Approvals, Registrations, and Patient Consent -----	12
2.1.2 Samples -----	12
2.2 Image Analysis -----	13
2.2.1 Study Devices and Settings -----	13
2.2.2 Morphometric Analyses and Quantifications-----	13
2.2.3 Digital Restoration Methods -----	14
2.3 Preliminary Statistics -----	17
2.3.1 Preliminary Estimation between Different Staining Methods-----	17
2.3.2 Correlation Analyses -----	18
2.3.3 Multiple Regression Analyses -----	18
2.3.4 Residual (Predictive Error) Analysis-----	18
2.3.5 Initial Age Grouping -----	19
2.3.6 Tests for Normality and Variance-----	19
2.3.7 Comparisons among Age Groups (1–6, 6–7, and 7–11 years) -----	19
2.3.8 Effect Sizes -----	19
2.3.9 Logistic Regression Analysis -----	20
2.4 Main Statistics and Analytical Framework-----	20
2.4.1 Segmented Regression Analysis-----	20
2.4.2 Bayesian Statistical Analyses -----	21
2.4.3 Monte Carlo Simulations -----	22
2.4.4 Revised Variance Analysis -----	23
2.4.5 Statistical Significance and Reporting of Results-----	23
2.4.6 Data Availability -----	23
3. Results -----	24
3.1 Summary and Initial Quality Control -----	24
3.1.1 Patient Selection and Analysis Dataset -----	24
3.1.2 Comparison of the H&E and G-T Staining Methods-----	26

3.2 Histopathology -----	27
3.2.1 Histopathological Findings-----	27
3.2.2 Correlation between Histopathological Parameters and Age-----	29
3.3 Exploratory Analytic Approach -----	31
3.3.1 Multiple Regression Analysis and Predictive Error -----	31
3.3.2 Age-Related Variability in Histopathological Parameters -----	36
3.3.3 Age-Stratified MFD Distributions-----	36
3.3.4 MFD Thresholds in Logistic Regression Analysis-----	41
3.4 Confirmatory Statistical Analysis-----	42
3.4.1 Segmented Regression Analysis-----	42
3.4.2 MFD Replacement by ATPase-Stained Slide Samples-----	44
3.4.3 Bayesian Analyses and Monte Carlo Simulations -----	46
3.4.4 Revised Variance Analysis -----	50
4. Discussion-----	54
4.1 Positioning of the Current Findings and Image-Analytic Innovations -----	54
4.1.1 Overview-----	54
4.1.2 Morphometric Methodologies and Image Analyses -----	54
4.2 From Exploratory to Confirmatory: Statistical Trajectory and Rationale-----	56
4.2.1 Exploratory Statistical Approaches -----	56
4.2.2 Confirmatory Statistics -----	59
4.2.3 Suggested Histopathological Inflection Point -----	61
4.3 Clinical Insights and Context of Use-----	62
4.3.1 Clinical Relevance-----	62
4.3.2 Implications for Disease Progression Assessment -----	64
4.3.3 Theoretical Clinical–Statistical Integration for Practice -----	67
4.4 Limitations and Future Directions -----	68
4.4.1 Study Limitations-----	68
4.4.2 Issues of This Study and Perspectives for Future Research -----	68
5. Conclusions -----	72
Acknowledgements -----	73
Conflicts of Interest-----	74
References-----	75

List of Publications and Conference Presentations

1. Publication

The following paper, based on part of this study, has been published:

Yamakado T, Ishikawa Y, Ise K, Wang L, Oda Y, Tsuda M, Kimura T, Nagao M, Ito M, Tanaka S and Tanei Z-I. Myofibre Density Reveals a Critical Threshold around Age 6 in Steroid-Naïve Duchenne Muscular Dystrophy: A Retrospective Observational Study. *Neuropathol Appl Neurobiol.* 2026 Apr;52(2):e70068. doi: 10.1111/nan.70068. PMID: 41848212; PMCID: PMC12997517.

2. Conference Presentation

Part of this study has been presented at the following conferences:

- Tanei Z-I, Yamakado T, Oda Y, Wang L, Tsuda M and Tanaka S. Autopsy and research of the nervous system at Hokkaido University. The 114th Annual Meeting of the Japanese Society of Pathology, April 17, 2025, Sendai, Japan.
- Yamakado T, Ishikawa Y, Ise K, Wang L, Oda Y, Tsuda M, Nagao M, Ito M, Tanaka S, Tanei Z-I. Myofiber Density Reveals a Critical Period Around Age 6 in Duchenne Muscular Dystrophy. The 115th Annual Meeting of the Japanese Society of Pathology, April 18, 2026, Sapporo, Japan.

Abstract

Background and Objectives: Duchenne muscular dystrophy (DMD) is a severe X-linked neuromuscular disorder. Despite therapeutic advances significantly improving patient outcomes, critical gaps remain in our understanding of the early-stage DMD pathophysiology. Systematic investigations into early-stage histopathological changes remain limited. A previous study has proposed a critical histological period around ages 6–7, marked by rapid increases in connective tissue, implying relative histological stability before this interval. This hypothesis was questioned and met with skepticism due to the absence of corresponding physiological phenomena around this specific age interval and the significantly earlier onset of clinical symptoms (typically between ages 3 and 5 years). Recognizing this significant clinicopathological discrepancy, the current study intended to resolve this contradiction. The present study leveraged severely deteriorated archival muscle biopsy specimens obtained over 40 years ago, without accompanying clinical records, from the former Yakumo National Hospital (now integrated into National Hospital Organization Hokkaido Medical Center, Japan). This study aimed to establish novel histopathological markers, clarify early disease mechanisms, and inform optimal treatment timing in early-stage DMD based solely on histopathological findings and age at biopsy (*ABx*).

Materials and Methods: This retrospective observational study adhered to the Declaration of Helsinki and was approved by the institutional review board of Hokkaido University Hospital (No. 024-0310), which oversees National Hospital Organization Hokkaido Medical Center. The requirement for informed consent was waived due to the retrospective study design. Archival muscle biopsy slides (quadriceps femoris muscle) of steroid-naïve DMD patients diagnosed between 1973 and 1984 were reviewed. Due to limitations in clinical record availability, initial selection partially relied on historical recollections. A total of 46 biopsies were evaluated among three pathologists. Histological slides (frozen and formalin-fixed paraffin-embedded) stained with hematoxylin-eosin (H&E), Gomori trichrome (G-T), and Masson trichrome (M-T) underwent digital restoration using look-up tables (LUTs) and Focus Stacking techniques. Actomyosin adenosine triphosphatase (ATPase)-stained slides were not included in the primary analysis but used in the sensitivity analyses. Quantitative morphometric analyses involving myofiber size (MFS)-related metrics (e.g., standard deviation of MFS (*Sd*), coefficient of variation of MFS (*Cov*)), myofiber density

(*MFD*), myofiber area (*MFA*), and connective/fibrotic tissue area (*CFA*) were performed using Adobe Photoshop. Statistical analyses involved preliminary comparisons between staining methods, correlation analyses, and multiple regression modeling to predict *ABx*. Model fit was assessed by calculating Akaike's information criterion (AIC), and prediction accuracy was validated by calculating the root mean squared error (RMSE) through leave-one-out cross-validation (LOOCV). Model residuals were analyzed to identify prediction error patterns across the predefined age intervals (1–4, 1–6, 6–7, 7–9, 9–11, and 11–17 years) using the Kruskal–Wallis test. Group differences were tested using analysis of variance (ANOVA) or non-parametric methods depending on data distribution, along with effect size analysis. Logistic regression was performed to identify critical thresholds of the parameters of interest at the age of 6. Segmented regression was performed to determine the existence of an age-related inflection point in the dynamics of each parameter. Bayesian statistical analyses and Monte Carlo simulations were subsequently conducted to confirm robustness and estimate the sample size. The primary measure was defined as the *MFD* age-breakpoint. Secondary measures comprised statistical power of breakpoint detection, age-predictive *MFD* cut-offs, and cross-validation of conventional parameters across age bands defined by the breakpoint. A significance threshold of $p < 0.05$ was applied to all statistical tests.

Results: A preliminary paired analysis revealed no significant difference in *MFA* between H&E and G-T staining (Cliff's δ : 0.015 [confidence interval (CI) $-0.28, 0.31$]; negligible), with strong correlation confirming interchangeability (Pearson's $r = 0.970$, $p < 0.001$); comparison with M-T staining was precluded by limited sample size. Bayesian analysis also confirmed negligible batch effects. Subsequently, from 46 muscle biopsies, 38 samples (patients aged 1–16 years) met the inclusion criteria. Histopathological evaluation demonstrated marked reductions in *MFD* with increasing age, alongside characteristic DMD features including increased connective tissue, variability in MFS, fatty infiltration, and internally nucleated and/or hypercontracted fibers. Correlation analyses identified *MFD* as the strongest predictor inversely correlated with *ABx* ($\rho = -0.85$). Multiple regression established robust age-prediction models, predominantly driven by *MFD* and its interaction with *Sd* (adjusted $R^2 = 0.79$). This full-range model (all ages included) yielded an AIC of 53.84 and an RMSE of 0.49. The results of Kruskal–Wallis tests demonstrated that model accuracy significantly declined in older age groups ($p < 0.01$). Consequently, an alternative model limited to patients younger than 7 years was evaluated, revealing a markedly improved fit (AIC = 29.12) and

comparable RMSE (0.47). When stratified by age groups (1–6 years, $n = 13$; 6–7, $n = 8$; 7–11, $n = 14$), ANOVA revealed substantial differences in *MFD* between groups (Glass's $\Delta = -1.75$ [CI $-2.85, -1.25$] for ages 1–6 vs. 6–7; -1.81 [CI $-2.91, -1.35$] for 1–6 vs. 7–11). Other key parameters, such as MFS-related parameters, *MFA*, and *CFA*, also exhibited significant differences between age groups 1–6 and ≥ 7 years. Replacing *MFD* values with measurements from ATPase-stained slides further supported these findings. Logistic regression revealed statistically significant *MFD* thresholds of 546–572 fibers/mm² differentiating patients below or above age 6. Segmented regression identified a breakpoint at age 6.25 [CI 5.08, 7.42], with Bayesian analysis and Monte Carlo simulations further confirming statistical support for this threshold and sufficient statistical power, approximately 80% probability of detecting the breakpoint within ± 1.25 years. Revised variance analysis stratified by breakpoint CI confirmed cross-parameter validity.

Discussion: These findings challenge previous views by demonstrating a rapid and significant *MFD* decline before age 6, indicating a previously unrecognized phase of accelerated myofiber loss. Methodologically, employing archival biopsy specimens, innovative image restoration techniques were applied to recover faded histological details from archival biopsy specimens. Exploratory statistical analyses incorporating regression modeling and prediction error analyses revealed precise age-related histopathological shifts. An alternative regression model restricted to patients younger than approximately 7 years old predicted *ABx* more accurately, highlighting increased variability in disease progression beyond this age. Variance analysis demonstrated superior sensitivity of *MFD* compared to conventional histological indices. Exploratory logistic regression suggested that *MFD* thresholds could distinguish between patients younger and older than age 6. Segmented regression confirmed the breakpoint around age 6. These thresholds closely align with clinically observed functional declines around ages 6–7 years, reinforcing a critical therapeutic window. Thus, *MFD* could inform optimal timing for initiating corticosteroids and novel therapies. Development of noninvasive monitoring tools is anticipated in clinical management. Importantly, whether the reduced *MFD* after age 6 signifies irreversible muscle pathology or if therapeutic interventions can effectively reverse these changes remains the subject of continued investigation. Collectively, this study underscores the importance and inherent challenges of developing new evaluation frameworks outside established analytical paradigms, delineating novel histopathological hallmarks of early-stage DMD

and informing future therapeutic strategies.

Conclusions: This study established *MFD* as a sensitive biomarker for early-stage DMD, identifying a previously unrecognized phase of rapid myofiber loss before age 6. This threshold age may represent a critical histopathological inflection in disease progression. Future research should prioritize longitudinal, multicenter studies to validate the robustness of *MFD* dynamics. The author believes that the present findings promote the broader utilization of archival specimens in future translational research, refining the understanding of the DMD's pathophysiology and enhancing clinical decision-making for this devastating neuromuscular disorder.

List of Abbreviations

ABx (parameter)

age at biopsy

AIC

Akaike's Information Criterion

ANOVA

analysis of variance

ATPase

actomyosin adenosine triphosphatase

BF₁₀

Bayes Factor (alternative hypothesis "1" over null hypothesis "0")

BIC

Bayesian Information Criterion

CFA (parameter)

connective/fibrotic tissue area

CI_s

confidence intervals

CK

creatine kinase

COU

Context of Use

Cov (parameter)

the coefficient of variation of myofiber size

CrI

credible interval

DMD

Duchenne muscular dystrophy

***Fat* (parameter)**

fatty degeneration area

FDA

Food and Drug Administration (United States)

FFPE

formalin-fixed paraffin-embedded

FOIs

fields of interest

G-T

Gomori trichrome

H&E

Hematoxylin-Eosin

***IntN* (parameter)**

percentage of internally nucleated fibers relative to *MFD*

LED

light-emitting diode

LOESS

Locally Estimated Scatterplot Smoothing

LOOCV

leave-one-out cross validation

LOOIC

leave-one-out information criterion

LUTs

Look-Up Tables

***Mean* (parameter)**

mean MFS

***MFA* (parameter)**

myofiber area

***MFD* (parameter)**

myofiber density

MFS

myofiber size

M-T

Masson trichrome

***NFA* (parameter)**

necrotic fiber area

***Opaque* (parameter)**

percentage of opaque fibers relative to *MFD*

ORs

odds ratios

PCR

polymerase-chain reaction

***RFA* (parameter)**

regenerative fiber area

RMSE

root mean squared error

ROPE

Region of Practical Equivalence

SD

standard deviation

***Sd* (parameter)**

standard deviation of MFS

1. Introduction

Background

The inception of this research originated from the concept that historical muscle biopsy slides, especially those from patients with Duchenne muscular dystrophy (DMD) preserved at the former National Yakumo Hospital (currently integrated into the National Hospital Organization Hokkaido Medical Center, Sapporo, Hokkaido, Japan), could be valuable for further investigative purposes. DMD is a well-documented, severe, X-linked disease characterized by progressive deterioration of skeletal muscle, causing profound muscular weakness and ultimately leading to severe cardiopulmonary complications (Birnkrant et al., 2018a, 2018b, 2018c; Crisafulli et al., 2020; Duan et al., 2021).

Historically, respiratory failure was the leading cause of death in DMD patients; however, with advances in assisted ventilation, cardiac failure has become predominant (Birnkrant et al., 2018b; Nonaka and Nishino, 2021). Despite remarkable advancements in clinical interventions and management (Bez Batti Angulski et al., 2023; Biggar et al., 2022; Birnkrant et al., 2018a, 2018b, 2018c; Duan, 2018; Duan et al., 2021; Elangkovan and Dickson, 2021; Erkut and Yokota, 2022; Frank et al., 2020; Happi Mbakam et al., 2022; Matthews et al., 2016; Mendell et al., 2020a, 2013; Min et al., 2019; Sun et al., 2020; Takeda et al., 2021; van Dommelen et al., 2024), there are notable gaps in the current understanding of the pathophysiology of early-stage DMD, particularly regarding the establishment of evident histological markers of disease progression.

What is Known and Unknown

Briefly, although histological evaluation is recognized as essential for understanding DMD pathology within the field of muscle histopathology (Bell and Conen, 1967; Desguerre et al., 2009; Dubuisson et al., 2022; Peverelli et al., 2015), surprisingly little is known about the detailed histopathological changes occurring during the early stages of DMD. Therefore, thoroughly reviewing previously established histopathological findings before starting this study project was critical. However, systematic investigations into histopathological correlations in DMD remain limited, with most studies primarily assessing myofiber cross-sectional area and fibrotic changes (Cardone et al., 2023; Desguerre et al., 2009; Peverelli et al., 2015).

The study by Peverelli et al. exerted significant influence on the initial research

design, prompting investigators to adopt methodologies similar to their image analysis approach (Peverelli et al., 2015). Central to their findings was the hypothesis that histological muscle architecture remains largely unchanged until approximately 6–7 years of age.

Ever since it was proposed, this hypothesis has met profound skepticism stemming primarily from the lack of any known significant physiological events around 6–7 years of age. If any established physiological phenomena, such as a marked increase in muscle strength or growth spurts, coincided precisely with this age range, such a claim might have been more convincing. However, no such physiological basis has been reported. Additionally, clinical manifestations of DMD clearly emerge much earlier than the critical period proposed by Peverelli et al. For instance, early clinical manifestations, including stair-climbing difficulties, waddling gait, and frequent falls—typically emerge between 3 and 5 years of age (Birnkrant et al., 2018b; Crisafulli et al., 2020; van Dommelen et al., 2024), with contractures of the iliotibial bands, hip flexors, and heel cords becoming apparent before 6 years of age (Brooke et al., 1983). The conclusion was that existing histopathological findings at this time could not adequately explain the pathophysiology of DMD.

Purposes

Given this context, resolving the significant discrepancy, a severe "clinicopathological dissociation," became a crucial challenge for this study. The only resources available for this study were significantly deteriorated archival specimens prepared over 40 years ago, lacking any accompanying clinical information (e.g., assessments of motor and/or physical function, serum creatine kinase (CK) levels). Nevertheless, this investigation was performed under these constraints, relying solely on histopathological findings and biopsy age. This study aimed to establish novel histopathological markers, clarify early disease mechanisms, and elucidate clinically meaningful correlations in early-stage DMD, ultimately to inform the optimal timing for initiating treatments such as corticosteroids and emerging advanced therapies.

2. Materials and Methods

1. Patients

2.1.1 Standard Protocol Approvals, Registrations, and Patient Consent

All procedures performed in this retrospective observational study involving human participants were conducted in accordance with the ethical standards of the institutional and national research committees, as well as the 1964 Declaration of Helsinki and its later amendments. The study complied with the Ethical Guidelines for Medical and Biological Research Involving Human Subjects established by the Ministry of Education, Culture, Sports, Science and Technology; and the Ministry of Health, Labor, and Welfare, Japan. The study was approved by the central institutional review board of Hokkaido University Hospital (Approval No. 024-0310, Sapporo, Hokkaido, Japan), which also oversees the National Hospital Organization Hokkaido Medical Center as part of its centralized review process. Due to the retrospective nature of the study, the requirement for informed consent was waived by the review board.

2.1.2 Samples

Initially, archival muscle biopsy slides had been stored under suboptimal conditions at an off-site storage facility, and many clinical records were either unavailable or incomplete. Therefore, preliminary selection of biopsies was partly informed by recollections provided by nurses familiar with patient histories at the Yakumo National Hospital. Subsequently, the clinical records retrieved from the hospital archives were reviewed to confirm basic patient information, enabling the identification and screening of a total of 46 muscle biopsy slides obtained from the quadriceps femoris muscles of patients diagnosed with DMD between 1973 and 1984.

Owing to the inherent limitations imposed by the retrospective nature of this study and the condition of the archival specimens available, the inclusion criteria were determined as follows: (1) pathological and clinical diagnosis of DMD, with genetic confirmation when feasible after the genetic tests were introduced around 2000; (2) treatment history at the former Yakumo National Hospital (subsequently integrated into the National Hospital Organization Hokkaido Medical Center in 2020) from 1968 to 2025; (3) no prior steroid treatment at the time of biopsy; and (4) sufficient quantity of muscle tissue and acceptable quality of the slide, confirmed through evaluation involving the author and two other independent pathologists. Slides that did not satisfy these inclusion criteria were excluded from further analysis. Patients diagnosed with

Becker muscular dystrophy were also excluded from this investigation. Archival biopsy specimens consisted of either frozen sections, formalin-fixed paraffin-embedded (FFPE) blocks, or in some cases, both, depending on availability. The primary histological staining methods used were hematoxylin-eosin (H&E), Gomori trichrome (G-T), and Masson trichrome (M-T); however, not all samples were available for every staining method. Actomyosin adenosine triphosphatase (ATPase)-stained slides were excluded from the main analysis but were instead utilized specifically for validation of statistical results and for sensitivity analyses.

2.2 Image Analysis

2.2.1 Study Devices and Settings

As described later, since this study utilized a look-up table (LUT)-based digital restoration method, reproducibility was anticipated to be significantly affected by variations in color temperature. Thus, using a microscope equipped with a light-emitting diode (LED) illumination system was essential to ensure consistent results. Therefore, microphotographic imaging was performed using a Nikon ECLIPSE Ci microscope, which utilizes an LED illumination system, equipped with a Nikon PLAN APO λ D 40 $\times$ objective lens and a Nikon DS-Fi3 camera, in conjunction with Nikon NIS-D software (Nikon Corporation, Inc., Tokyo, Japan). Fields of interest (FOIs) for image analysis were selected by consensus of the three pathologists. High-resolution images were captured at 400 $\times$ magnification ($2,880 \times 2,048$ pixels; spatial resolution: $0.11 \mu\text{m}/\text{pixel}$) and subsequently assembled into large composite images, each comprising 16 individual tiled frames, resulting in a final image resolution of $11,520 \times 8,192$ pixels. Each composite image corresponded to an area of approximately 1.14 mm^2 as visualized through a 10 $\times$ objective lens. Semi-quantitative digital analysis of the resulting images was conducted using Adobe Photoshop 2024 (Adobe Systems, Inc., San Jose, CA, USA), a software platform extensively employed in histopathological studies (Ding et al., 2022; Egan et al., 2012; Fitzgerald et al., 2019; Zhou et al., 2017).

2.2.2 Morphometric Analyses and Quantifications

Although a semi-quantitative analytical approach involving differentiation between normal and abnormal myofibers (e.g., normal/total myofiber area measurements and normal/total myofiber counting) was initially adopted, subsequent evaluations raised concerns regarding the reproducibility, necessity, and validity of distinguishing myofiber states based on qualitative criteria. Consequently, the parameter

definitions were revised and finalized as described below, emphasizing fully quantitative and reproducible measures.

In this report, the variables used in the statistical analyses are emphasized by capitalizing the initial letter and italicizing (e.g., *Parameter*) to aid clarity. Four variables related to the myofiber size (MFS) were initially introduced:

- (1) The mean MFS (*Mean*, μm^2);
- (2) The standard deviation of MFS (*Sd*, μm^2);
- (3) The coefficient of variation of MFS (*Cov*); and
- (4) Myofiber density (*MFD*, fibers/ mm^2), which was derived as an auxiliary metric during the enumeration of individual fibers while evaluating MFS-related parameters, with an exclusion rule for truncated myofibers at the field edge, regardless of pathological features (e.g., central nucleation, hypercontraction).

In addition, seven conventional histological variables, established in previous research (Desguerre et al., 2009; Dubuisson et al., 2022; Peverelli et al., 2015), were quantified, each described as a percentage per FOI unless otherwise specified:

- (1) Myofiber area (*MFA*), similarly regardless of pathological features;
- (2) Connective/fibrotic tissue area (*CFA*);
- (3) Necrotic fiber area (*NFA*);
- (4) Regenerative fiber area (*RFA*);
- (5) Fatty degeneration area (*Fat*);
- (6) Percentage of opaque fibers relative to *MFD* (*Opaque*, count/*MFD*, %); and
- (7) Percentage of internally nucleated fibers relative to *MFD* (*IntN*, count/*MFD*, %).

2.2.3 Digital Restoration Methods

Color restoration was performed utilizing LUTs within the Nikon NIS-D software, an approach commonly used in fluorescence and confocal microscopy (Li et al., 2011; Sanderson et al., 2014). Real-time high-resolution images captured at 400 $\times$ magnification with LUT adjustments were subsequently combined into large composite images (**Figure 1A**). The application of LUTs significantly improved visualization, facilitating the identification of *IntN* and *Opaque*, as well as clearly delineating myofiber borders, fatty infiltration, and other histological features (**Figure 1B**). To minimize artifacts due to focal-plane discrepancies, a Focus Stacking (Jütte et al., 2022) was implemented using a script within Adobe Photoshop. This approach, originally inspired by photographic techniques aimed at achieving uniform focus across multiple

focal planes (Shosekai-Kenkyujo, 2021), combines several images obtained at different depths into a single composite image with extended depth-of-field clarity (**Figure 1C**). Regions exhibiting residual artifacts that could not be adequately corrected by these methods were excluded from further analysis. Individual myofibers were traced using the Magnetic Lasso Tool in Adobe Photoshop (contrast: 40, frequency: 100). To further enhance contrast and measurement accuracy, images were converted to monochrome before assessing MFS-related parameters (**Figure 1D**).

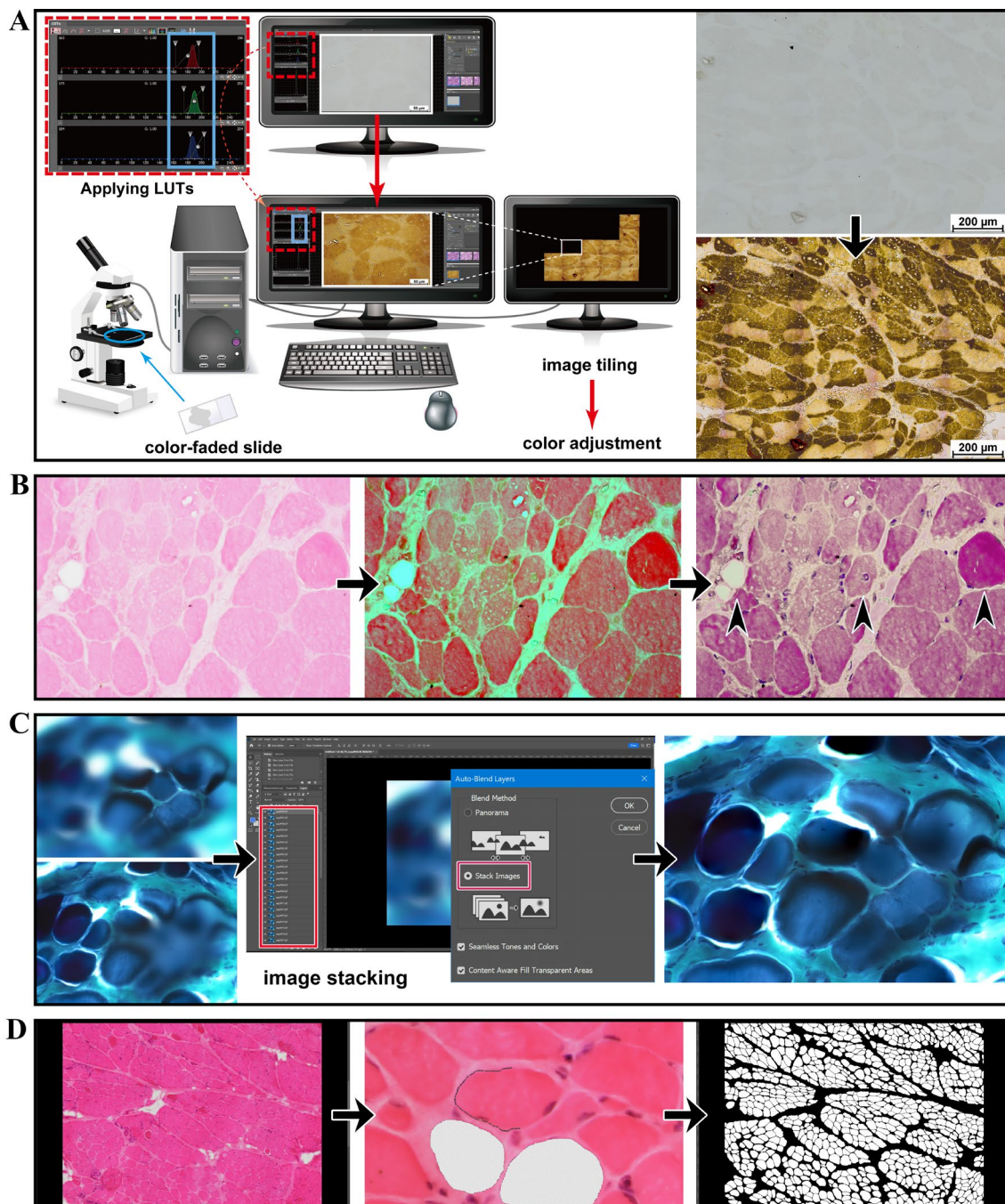


Figure 1: Workflow for image restoration and semi-quantitative morphometric analysis

A: This panel demonstrates the digital restoration of an extensively faded ATPase-stained specimen (pH 9.4), representing one of the most severely deteriorated examples within this study. Real-time adjustment of LUTs in the live camera preview mode substantially enhanced the visibility of tissue structures obscured by fading. Individual restored images were subsequently tiled together to create a high-resolution composite image covering the entire FOI. Additional standard image-processing tools in Photoshop, such as Color Balance and Hue/Saturation adjustments, were then used to restore a characteristic ATPase-like appearance. This LUT-based restoration approach, initially optimized for ATPase staining, served as a foundation for refining methods applicable to other histological stains illustrated in subsequent panels.

B: The restoration procedure detailed in panel A was adapted here for severely faded H&E-stained specimens, in which nuclei were particularly difficult to visualize. Initially, LUT adjustments rendered previously invisible nuclear structures green, significantly improving visualization of nuclear morphology and myofiber boundaries. Subsequent color adjustment in Photoshop transformed the highlighted green nuclei into their typical purple color, effectively restoring the appearance characteristic of H&E staining. This method greatly enhanced the identification of centrally nucleated fibers, hypercontracted fibers, and regions of fatty infiltration (indicated by arrowheads). Panels A and B thus highlight the common strategy of employing LUT-based digital enhancement followed by targeted color adjustments to recover faded histological details.

C: This panel illustrates the application of Focus Stacking to enhance visualization of a G-T-stained specimen. Multiple images captured at different focal depths were combined using the Auto-Blend Layers function in Adobe Photoshop, resulting in a uniformly focused, high-resolution composite image. This technique successfully preserved intricate structural details of myofibers, significantly improving overall image quality and accuracy.

D: This panel demonstrates how individual myofibers were delineated in an H&E-stained specimen using the Magnetic Lasso Tool in Adobe Photoshop (contrast: 40, frequency: 100). Prior to fiber tracing, a subtle LUT-based enhancement was applied to improve image contrast slightly, facilitating accurate boundary delineation. After segmentation, the resulting image was converted to monochrome to simplify visual interpretation and enable precise quantitative analyses. Subsequently, myofiber size

(initially measured in px^2) and associated standard deviation values were calculated using Microsoft Excel (Microsoft Corporation, Redmond, WA, USA).

Note: Panel A was created using Adobe Illustrator software (Adobe Systems, Inc., San Jose, CA, USA) and includes freely available resources from Pixabay

(<https://pixabay.com>) combined with screenshots from Nikon NIS-D software (Nikon Corporation, Inc., Tokyo, Japan). Panels C and D feature screenshots from Adobe Photoshop, which served as the primary image analysis software used in this workflow.

(Abbreviations: LUTs, look-up table; ATPase, actomyosin adenosine triphosphatase; FOI, field of interest; H&E, hematoxylin-eosin; G-T, Gomori trichrome; *MFA*, myofiber area; *MFS*, myofiber size.)

Individual myofiber areas were measured using the Analysis tool in Adobe Photoshop, with each myofiber carefully traced and separately delineated to minimize measurement inaccuracies due to overlapping or adjacent fibers. The software automatically assigned unique labels to each outlined myofiber, facilitating subsequent calculation of morphometric parameters as previously described. Data obtained from these morphometric analyses were exported from Adobe Photoshop as text files, subsequently imported into Microsoft Excel for further processing, and finally saved in the CSV format for use in statistical analyses.

2.3 Preliminary Statistics

All statistical analyses were performed using R (version 4.3.2). Details on the availability of raw data and analysis scripts are provided later. The analytical procedures are outlined as follows:

2.3.1 Preliminary Estimation between Different Staining Methods

Two frozen staining methods were compared: H&E and G-T staining. As described, due to limitations in sample availability, this study included H&E-, G-T-, and M-T-stained slides. To verify the comparability between these staining methods, *MFA* values obtained from H&E and G-T methods were evaluated using paired analyses. First, a small preliminary sample set ($n = 16$) was used in the Wilcoxon signed-rank tests and power analysis to estimate the required sample size for achieving adequate statistical power. Subsequently, the evaluation was expanded to a larger paired-sample set, and the Wilcoxon signed-rank test was performed. Additionally, Pearson's product-moment correlation coefficients were calculated to further validate methodological

interchangeability between the two staining methods. Due to limited slide availability, comparative analyses involving M-T staining, as well as comparisons between frozen and FFPE sections, could not be performed. Furthermore, the practical equivalence of H&E and G-T staining methods was quantified by a Bayesian paired-sample t-test (Jeffreys–Zellner–Siow Cauchy prior with scale parameter $r = 0.707$) (Rouder et al., 2009). A Region of Practical Equivalence (ROPE) of $|d| < 0.30$ was applied (Lakens, 2022); practical equivalence was concluded when $\geq 80\%$ of the posterior mass fell inside the ROPE and Bayes Factor (BF_{10}) < 1 . Additionally, a linear mixed-effects model with the paired observations clustered by specimen subject number was fitted, and changes in the specimen-level variance component (τ^2) were examined when staining method was added as a fixed effect in the model.

2.3.2 Correlation Analyses

Bivariate correlation analyses using Spearman's rank correlation coefficients were performed to assess relationships between each quantified parameter and age at biopsy (*ABx*).

2.3.3 Multiple Regression Analyses

All the variables were treated as continuous and transformed into logarithmic scales followed by centering and/or z-score standardization. Multiple regression analysis with stepwise selection was performed to identify significant predictors for similarly transformed *ABx*. Potential interaction terms between variables were also explored. Akaike's information criterion (AIC) and model diagnostics, including residual plots, QQ plots, and Cook's distance, were examined to ensure the validity of the assumptions and reliability of the regression models. Additionally, the robustness and predictive reliability of the established regression models were evaluated by calculating the root mean squared error (RMSE) and standard deviation (SD) using leave-one-out cross-validation (LOOCV).

2.3.4 Residual (Predictive Error) Analysis

An original residual analysis approach was developed and applied based on prediction errors (residuals) derived from the multiple regression model. Specifically, the distribution of prediction errors was compared across the following predefined age intervals: 1–4, 4–6 (or 1–6), 6–7, 7–9, 9–11 (or 7–11), and 11–17 years, using the Kruskal–Wallis test. The rationale for these detailed age groupings was based on three

primary considerations: (1) clinical symptoms typically manifest before age 6 years (Birnkrant et al., 2018b); (2) previous studies have hypothesized that a critical disease progression phase occurs around ages 6–7 years (Peverelli et al., 2015); and (3) to achieve a relatively balanced distribution of sample sizes within each age group, with the exception of the oldest group (11–17 years). Furthermore, additional validation was performed using RMSE values obtained through LOOCV to verify the stability and predictive accuracy of the regression model for the younger age groups (up to 7 years).

2.3.5 Initial Age Grouping

Based on the results described above, participants were subsequently stratified into the following age groups: 1–6 years, 6–7 years, and 7–11 years. The 11–17 year age group was excluded due to the small sample size ($n = 3$; see the Results section).

2.3.6 Tests for Normality and Variance

The Shapiro–Wilk test was used to evaluate data normality, while Levene's test assessed the homogeneity of variances for each variable.

2.3.7 Comparisons among Age Groups (1–6, 6–7, and 7–11 years)

For normally distributed data with homogeneous variances, analysis of variance (ANOVA) was performed, followed by Tukey's HSD post-hoc test. For non-normally distributed data, the Kruskal–Wallis test was applied, with post-hoc comparisons performed using the Conover test, incorporating Bonferroni correction. Welch's ANOVA followed by the Games–Howell post-hoc test was used in cases where normality was assumed but variance homogeneity criteria were not met.

2.3.8 Effect Sizes

Effect sizes of comparisons between groups were calculated based on data distribution characteristics. For variables satisfying assumptions of normality and homogeneity of variance, Hedges' d with bootstrapped confidence intervals (CIs) (5,000 iterations) was calculated based on the approach described by Nakagawa and Cuthill (Nakagawa and Cuthill, 2007). For variables meeting normality but violating the homogeneity of variance assumption, Glass's Δ with bootstrapped CIs was calculated. For variables that violated the assumption of normality, Cliff's δ with bootstrapped CIs was employed. Additional analyses were conducted by excluding extreme outliers to assess the robustness of statistical findings, where appropriate. Additionally, post-hoc

power analysis was conducted to confirm whether the sample sizes provided sufficient statistical power (target statistical power: 80%).

2.3.9 Logistic Regression Analysis

For patients aged <11 years, logistic regression was employed to evaluate the relationship between the values of the parameter of interest and ABx , using 6 years as the reference point. The model fit was assessed using AIC, and predictive accuracy was assessed through odds ratios (ORs) with corresponding 95% CIs. Bonferroni correction for multiple comparisons was applied when interpreting the statistical significance. Additional analyses were conducted using alternative reference points at ages 4, 5, and 7 years.

2.4 Main Statistics and Analytical Framework

2.4.1 Segmented Regression Analysis

To objectively determine and validate the existence of a critical age-related inflection point in the dynamics of each parameter, segmented (piece-wise) regression analysis was performed following the method proposed by Muggeo (Muggeo, 2003). According to Muggeo, the initial breakpoint can be visually determined based on the inspection of plotted data (Muggeo, 2003); thus, in this study, a moving-average plot was used to estimate a candidate breakpoint. First, a simple linear regression model was constructed with biopsy age (x_i) as the predictor variable and each parameter as the response variable (y_i), restricting the analysis to patients younger than 11 years old ($ABx < 11$ years, $n = 35$) due to insufficient sample size in older age intervals ($n = 3$; see Results section). Next, segmented regression analysis was performed according to the following model:

Model:

$$y_i = \alpha + \beta_1(x_i - \psi)_- + \beta_2(x_i - \psi)_+ + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

$$\Leftrightarrow y_i = \beta_0 + \beta_1 x_i + \beta_\Delta(x_i - \psi)_+ + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

In this model, α denotes the intercept at the breakpoint, β_0 represents the intercept, β_1 and β_2 indicate slopes before and after the breakpoint ($\beta_\Delta = \beta_2 - \beta_1$), respectively, and ψ denotes the breakpoint (inflection point). The error term (ε_i) was assumed to be normally distributed with a mean of zero and constant variance ($\varepsilon_i \sim N(0, \sigma^2)$). Model improvement introduced by adding the breakpoint was statistically tested

using ANOVA, comparing the segmented regression model with the simple linear regression model.

2.4.2 Bayesian Statistical Analyses

Two Bayesian segmented regression analyses were performed. The first Bayesian segmented regression was performed using conjugate posterior distribution with the fixed breakpoint according to the general framework described by Gelman et al. as below (Gelman et al., 2025). In this analysis, the same segmented regression model was used:

Matrix form:

$$y | \beta, \sigma, X \sim N(X\beta, \sigma^2 I), \quad X = [1, x, (x - \psi)_+], \quad \beta = (\beta_0, \beta_1, \beta_\Delta)^T$$

Prior (noninformative):

$$p(\beta, \sigma^2 | X) \propto \sigma^{-2}$$

Conjugate posterior:

$$\begin{aligned} \hat{\beta} &= (X^T X)^{-1} X^T y, \quad S = (y - X\hat{\beta})^T (y - X\hat{\beta}), \quad s^2 = S/(n - k) \\ \sigma^2 | y &\sim \text{Inv-}\chi^2(n - k, s^2), \quad \beta | \sigma^2, y \sim N(\hat{\beta}, \sigma^2 (X^T X)^{-1}) \end{aligned}$$

Under the segmented regression model, matrix form, prior, and conjugate posterior defined above, posterior samples of the regression coefficients were obtained by sampling from the posterior multivariate normal distribution, for which the covariance matrix $(X^T X)^{-1}$ was decomposed via Cholesky decomposition and subsequently multiplied by standard normal random vectors, ensuring stable and computationally efficient posterior sampling (Gelman et al., 2025). The model fit was assessed by calculating the Bayesian Information Criterion (BIC) using Schwarz's approximation (Kass and Wasserman, 1995; Volinsky and Raftery, 2000). The ROPE was empirically defined based on measurement errors derived from paired observations between original and ATPase-stained values of the parameter of interest. Furthermore, an additional sensitivity analysis using expanded ROPE ranges (ROPE +1SD, +1.5SD, +2SD) was performed to further examine the stability of the findings.

The second Bayesian segmented regression analysis was conducted using Markov Chain Monte Carlo methods. In this Bayesian model specification, ABx centered at the mean (x_i^*) was modelled as the predictor variable, and the parameter of

interest (y_i) as the response variable, defined as follows:

Model:

$$y_i = \beta_0 + \beta_1(x_i^* - \psi^*) + \beta_\Delta(x_i^* - \psi^*) \cdot \text{step}(x_i^* - \psi^*) + \varepsilon_i, \quad \varepsilon_i \sim N(0, \sigma^2)$$

$$\mu x = \text{mean } ABx (<11 \text{ years}), \quad x_i^* = x_i - \mu x, \quad \psi^* = \psi - \mu x$$

Following Gelman et al., we specified a normal linear regression ($y \sim N(X\beta, \sigma^2 I)$) with weakly informative priors and obtained posterior draws via the Hamiltonian Monte Carlo algorithm with the No-U-Turn Sampler (Gelman et al., 2025). Weakly informative priors were selected based on estimates from the frequentist segmented regression analysis. β_0 denotes the intercept as well. A Bayesian sensitivity analysis was also performed using a ROPE to evaluate whether changes in slopes (β_1, β_Δ) at the estimated breakpoint (ψ^* , defined at zero on a centered scale) represented statistically significant differences. The model fit and predictive performance were compared with those of a simpler linear regression model using the leave-one-out information criterion (LOOIC).

2.4.3 Monte Carlo Simulations

Using the LOOCV-derived RMSE as the noise parameter, a Monte Carlo simulation with 2×10^6 draws (chosen so that the binomial standard error for misclassification probabilities ≤ 0.05 falls below 0.02 percentage points: Monte Carlo Error) converted analytic error into clinical risk by estimating the probability that a histological parameter-based rule would mis-classify patients as younger or older than 6.25 years across hypothetical age gaps of 0.25, 0.50, 0.75, 1.00, 1.25, 1.50, 1.75, and 2.00 years. Moreover, sample-size simulations using the Monte Carlo method were conducted to evaluate statistical power for breakpoint detection. Synthetic datasets were generated by randomly sampling age values ($ABx < 11$ years) from observed data with replacement. For each sampled dataset, response values of the parameter of interest were simulated based on estimated segmented regression variables (breakpoint, slopes, and RMSE). The simulations were repeated 10,000 times for each combination of sample size (20 to 70 in increments of 5) and acceptable error margins (± 1.0 , ± 1.25 , and ± 1.5 years). Breakpoint estimation precision was defined as the proportion of successful estimations (estimated breakpoint within the acceptable margin). Additionally, the second logistic model was inverted to locate the values of the histological parameters giving posterior probabilities of 0.80, 0.50, and 0.20 for being < 6.25 years; these cut-

points defined the high-, grey-, and low-probability zones that were summarized in a three-level decision table.

2.4.4 Revised Variance Analysis

To validate the behavior of histological parameters except for *MFD*, a group-wise variance analysis was performed. Age stratifications were defined based on the 95% CI of the breakpoint of *MFD*: Early (<CI-low), Transitional (CI span), and Late ($\geq$ CI-high). Normality and homogeneity were confirmed using the Shapiro–Wilk and Levene tests, respectively. These findings directed either one-way ANOVA with Tukey's post-hoc test, or the Kruskal–Wallis test with Conover post-hoc adjustment.

2.4.5 Statistical Significance and Reporting of Results

A significance threshold of $p < 0.05$ was applied consistently across all statistical tests. However, recent statistical reporting guidelines caution against relying solely on p -values as indicators of statistical significance (Harrington et al., 2019). Therefore, interpretation of statistical findings in the present study was based not exclusively on p -values, but also on associated measures such as effect sizes and corresponding 95% CIs. While these were particularly emphasized to quantify the magnitude and clinical relevance of the differences, statistical significance in these multiple comparisons was determined based on false discovery rate-adjusted p -values ($q < 0.05$), regardless of whether the corresponding 95% CIs crossed zero. Additionally, unless stated otherwise, the term "CI" refers explicitly to a 95% CI.

For clarity, given the complexity of the statistical workflow in this study, each Results subsection begins with a brief statement of the rationale for the subsequent analysis in light of the preceding findings, as a short guide for the readers.

2.4.6 Data Availability

All datasets and R scripts are publicly available at the following GitHub repository: <https://github.com/tyamakadopatho2-ops/Pathologist-Repository-for-Open-Science/tree/Doctoral-Dissertation>

3. Results

3.1 Summary and Initial Quality Control

3.1.1 Patient Selection and Analysis Dataset

From an initial number of 46 muscle biopsy slides, eight slides were excluded due to severe fading or insufficient tissue quantity, yielding data from 38 patients aged 1–16 years (**Table 1**); the staining methods were not standardized (**Table 2**).

Table 1: Summary of parameters with biopsy year and genotype information

subject	ABx ⁺	Biopsy year	Mean	Cov	Sd	MFD	MFA	Fat	CFA	Genotype†
1	1.0	1983	457.9	0.58	265.7	1471.9	68.9%	0.2%	29.6%	NA
2	1.9	1977	565.3	0.48	271.9	1227.0	72.3%	0.1%	26.7%	NA
3	2.0	1975	622.1	0.53	328.7	1069.5	68.7%	0.2%	30.7%	exon44–47 del
4	2.4	1979	598.1	0.56	336.2	1149.8	71.6%	1.7%	25.8%	exon56 del
5	3.3	1977	851.2	0.54	455.7	656.8	57.3%	0.7%	39.2%	-
6	3.7	1975	531.3	0.44	230.9	953.5	53.7%	0.2%	45.7%	exon44–47 del
7	4.0	1984	1037.9	0.53	551.2	563.2	53.5%	1.3%	41.6%	-
8	4.1	1979	577.5	1.20	693.4	867.2	53.4%	5.3%	40.3%	exon 44–51 del
9	4.3	1973	828.5	0.50	417.0	766.0	65.7%	0.6%	32.1%	exon56 del
10	4.6	1984	870.4	0.65	568.3	545.5	45.9%	1.7%	48.9%	exon44 del
11	4.7	1978	721.1	0.51	365.7	815.2	62.1%	0.8%	36.7%	exon36 del
12	5.0	1980	1007.4	0.50	503.6	571.1	59.5%	1.7%	37.3%	exon43–50 del
13	5.2	1980	624.3	0.53	330.3	976.4	64.9%	0.1%	34.0%	-
14	6.0	1975	725.1	0.82	591.6	552.0	38.5%	1.4%	58.7%	-
15	6.2	1977	1352.0	0.43	585.0	454.7	65.8%	0.1%	31.6%	-
16	6.2	1977	1176.2	1.11	1309.4	365.8	42.2%	1.4%	54.6%	NA
17	6.2	1978	556.3	0.88	489.5	559.9	29.2%	0.7%	67.1%	exon20, exon59 Mut
18	6.3	1979	1602.3	0.78	1244.7	300.7	51.7%	3.5%	44.4%	NA
19	6.4	1979	1520.6	0.95	1442.2	343.8	54.7%	3.1%	40.7%	NA
20	6.5	1979	1554.0	0.61	954.3	400.7	66.8%	0.5%	31.7%	exon6–8 del
21	6.6	1983	1683.7	0.70	1176.0	216.6	36.3%	1.5%	59.9%	-
22	7.0	1978	850.3	0.73	623.7	553.8	47.1%	3.9%	45.3%	-
23	7.2	1979	680.0	0.71	479.3	572.7	39.8%	13.2%	46.6%	exon46–48 del
24	7.2	1976	970.2	0.98	954.0	445.2	45.8%	3.6%	50.2%	NA
25	7.8	1980	772.0	0.60	464.8	458.1	36.4%	6.1%	57.0%	NA
26	8.3	1980	1231.8	1.72	2118.5	245.4	30.4%	6.4%	62.4%	NA
27	8.7	1982	1048.4	1.26	1324.3	497.9	53.9%	0.05%	41.7%	exon50 del
28	8.8	1981	1327.3	0.82	1081.8	282.9	38.6%	3.5%	55.0%	exon3 del
29	8.9	1979	613.6	0.86	526.5	386.5	23.5%	18.6%	55.9%	-
30	9.1	1982	1285.4	0.95	1215.0	423.0	59.0%	5.0%	33.6%	exon44 del
31	9.7	1976	889.0	0.75	666.2	291.0	25.4%	3.2%	68.2%	-
32	9.8	1982	1203.7	0.78	943.5	414.1	52.8%	8.6%	37.3%	exon51–60 del
33	10.3	1980	3643.5	0.99	3603.9	143.9	58.2%	10.3%	30.8%	exon12 dupl
34	10.7	1982	909.6	0.90	821.4	240.5	20.8%	6.0%	71.6%	exon44 del
35	10.9	1977	1565.2	0.77	1199.8	393.5	67.7%	3.0%	28.7%	exon56 del
36	12.8	1983	1028.5	1.07	1096.0	219.3	20.9%	3.4%	73.7%	NA
37	15.1	1973	1953.9	0.92	1787.6	170.2	33.9%	9.4%	54.6%	NA
38	16.3	1973	1504.0	1.46	2189.0	172.1	26.5%	15.4%	55.9%	exon46–49 del

(**Abbreviations:** *ABx*, age at biopsy; *Mean*, mean myofiber size; *Sd*, standard deviation of myofiber size; *Cov*, coefficient of variation of myofiber size; *MFD*, myofiber density; *MFA*, myofiber area; *Fat*, fatty degeneration area; *CFA*, connective/fibrotic tissue area; *del*, deletion; *Mut*, mutation; *dupl*, duplication; *MLPA*, multiplex ligation-dependent probe amplification; *PCR*, polymerase-chain reaction; *NA*, not available.)

* *ABx* was analyzed as a continuous parameter in correlation analyses and presented to one decimal place. To ensure analytical precision, the exact age was calculated by adding the fractional value of biopsy month (month divided by 12) to the *ABx*. For patients sharing the same age and month, rankings were further differentiated based on the week of biopsy.

† Genetic mutations were determined by MLPA or PCR testing. "NA" indicates cases where genetic testing was not performed, while a hyphen (-) signifies that no mutation was detected despite testing.

Table 2: Availability of samples for semi-quantitative image analysis

Subject	Frozen H&E	Frozen G-T	FFPE H&E	FFPE M-T	Frozen ATPase
1	○	○	×	×	○
2	○	○	×	×	×
3	×	×	○	○	×
4	○	○	○	×	○
5	○	○	×	×	×
6	×	×	○	○	×
7	○	○	×	×	○
8	○	○	×	×	×
9	○	○	○	○	×
10	○	○	×	×	×
11	○	○	○	○	×
12	○	○	×	×	○
13	○	○	×	×	○
14	×	×	×	○	×
15	○	○	×	×	×
16	○	○	×	○	×
17	×	×	○	×	×
18	○	○	×	×	×
19	○	○	×	×	○
20	○	○	×	×	○
21	○	○	×	×	○
22	○	○	×	×	×
23	○	○	×	×	×
24	○	×	×	×	×
25	○	○	×	×	×
26	○	○	×	×	×
27	○	○	×	×	×
28	○	○	○	○	×
29	○	○	×	×	○
30	×	○	×	×	○
31	○	○	×	×	×
32	○	○	×	×	○
33	○	○	×	×	○
34	×	○	×	×	○
35	○	○	×	×	×
36	○	○	×	×	×
37	×	○	×	×	×
38	×	○	×	×	○

(**Abbreviations:** FFPE, formalin-fixed paraffin-embedded; H&E, hematoxylin-eosin; G-T, Gomori Trichrome; M-T, Masson Trichrome; ATPase, actomyosin adenosine triphosphatase.)

○: available; ×: severe fading, insufficient tissue quantity, or loss.

3.1.2 Comparison of the H&E and G-T Staining Methods

At the time of this comparison, specimen collection and validation were still at an early stage, and the full set of biopsy samples had not yet been assembled. Consequently, to address methodological concerns regarding staining variability, a preliminary evaluation was performed on a limited subset of available paired specimens to determine whether different staining methods (H&E and G-T) could be reliably integrated in subsequent quantitative analyses. Based on this, the preliminary analysis of 16 paired specimens revealed no significant difference between *MFA* values obtained from H&E and G-T staining specimens (Wilcoxon signed-rank test, $p > 0.05$). Power analysis indicated that approximately 28 pairs would yield sufficient statistical power (80%). Thus, the sample size was increased to 28 paired specimens. Once again, no significant differences were found between *MFA* measurements obtained from the H&E- and G-T-stained slides (Wilcoxon signed-rank test, $p \approx 0.9$; Cliff's $\delta = 0.015$ [CI $-0.28, 0.31$], negligible). Subsequent power analysis revealed that detecting any significant differences would require an impractically large sample size (approximately $n = 450$), further showing negligible practical differences between the two staining methods. Additionally, Pearson's product-moment correlation coefficient demonstrated an extremely strong correlation between *MFA* values obtained from H&E and G-T staining methods ($r = 0.97$, $p < 0.001$; **Figure 2**), confirming statistical interchangeability based on the interpretation described by Bell and Conen (Bell and Conen, 1967). Validation using *CFA* was omitted because *MFA* and *CFA* showed a statistically significant and extremely strong correlation in the main correlation analysis (Pearson's $r = 0.96$).

Bayesian analysis showed that approximately 82% of the posterior for the standardized difference was within the ROPE (posterior Cohen's $d = 0.12$, 95% Highest Density Interval = -0.26 to 0.49). The BF_{10} favored the null hypothesis ($BF_{10} = 0.101$), showing substantial (almost strong) evidence according to the Jeffreys' criteria (Kass and Raftery, 1995). Mixed-effects modeling further demonstrated negligible specimen-level variance attributable to staining differences (variance reduction = 0.03%).

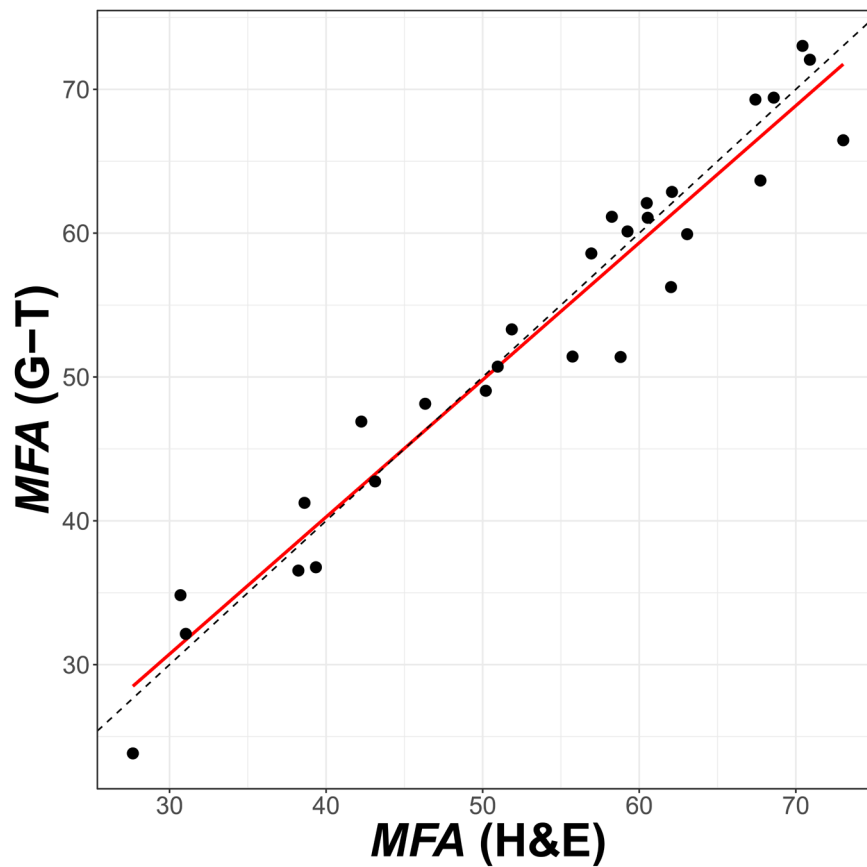


Figure 2: Correlation of *MFA* between H&E- and G-T-stained samples

The horizontal and vertical axes represent *MFA* values (%) from H&E- and G-T-stained slides, respectively. Pearson's product-moment correlation coefficient (r) between *MFA* values obtained from 28 pairs of matched H&E- and G-T-stained slides was approximately 0.97.

(**Abbreviations:** *MFA*, myofiber area; H&E, hematoxylin-eosin; G-T, Gomori trichrome.)

3.2 Histopathology

3.2.1 Histopathological Findings

Histological evaluation revealed distinct differences in *MFD* patterns between specimens obtained from patients younger than 6 years and those from patients in older age groups, as described in subsequent sections. A majority of samples exhibited histopathological alterations, including *MFS* variation, the presence of hypercontracted myofibers and internally nucleated myofibers, increase in connective tissue, reduced *MFA*, widened inter-myofiber spacing, and adipose tissue replacement (**Figure 3**).

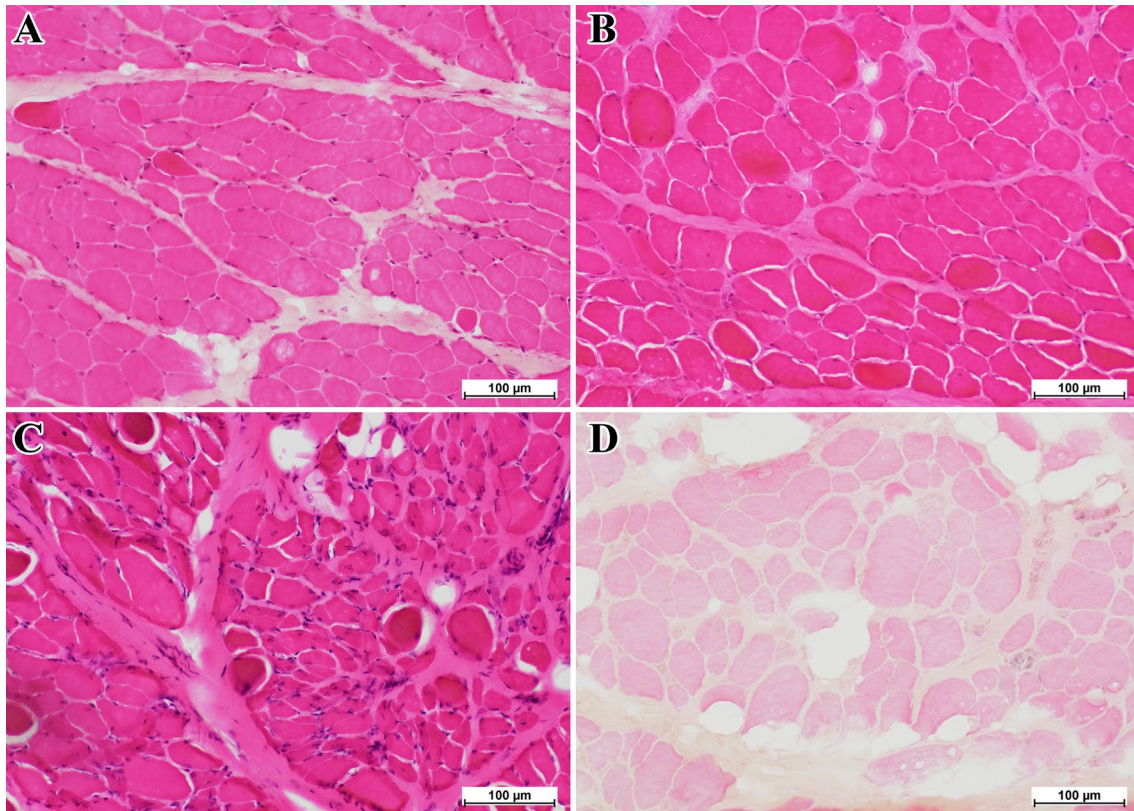


Figure 3: Representative H&E-stained images of DMD patients at different ages

Each panel shows representative microphotographs of H&E-stained specimens from patients of specific ages (each at 200× magnification): 4 years and 4 months (A), 6 years and 2 months (B), 7 years and 0 months (C), and 9 years and 9 months (D). Note that the original FOI was captured with a 10× objective lens, although the panels are displayed at 20× magnification.

A: In the 4-year-old patient, myofibers show consistent histological characteristics, exhibiting well-preserved *MFA* with an expanded connective tissue area.

B: In the 6-year-old patient, myofibers similarly show consistent morphology and *MFA* comparable to panel A; however, quantitative assessment reveals a substantial decrease in *MFD*. Comparing panels A and B, *MFA* values are nearly identical at 65.69% and 65.77%, respectively, while *MFD* declines notably from 765.96 to 454.67. The connective tissue area remains relatively unchanged, with *CFA* values of 32.06% (panel A) and 31.55% (panel B).

C: The 7-year-old patient demonstrates increased variability in *MFS*, accompanied by notably expanded connective tissue and initial signs of mild fatty infiltration.

D: The 9-year-old patient shows marked reductions in both *MFD* and *MFA*, alongside a pronounced expansion of connective tissue and distinct areas of fat replacement.

(**Abbreviations:** DMD, Duchenne muscular dystrophy; H&E, hematoxylin-eosin; FOI, field of interest; LUTs, look-up tables; *MFA*, myofiber area; *MFD*, myofiber density; *CFA*, connective/fibrotic tissue area; MFS, myofiber size.)

3.2.2 Correlation between Histopathological Parameters and Age

To identify new pathological metrics, correlations between the *ABx* and key histological parameters were assessed. All the selected variables contained complete datasets. The results of semi-quantitative image analyses are described in **Table 1**, along with biopsy year and genotype data. Among all parameters, *MFD* exhibited the strongest correlation with *ABx*, demonstrating a marked decline between 1 and 6 years (**Figure 4**). The Spearman's correlation coefficients for key variables were as follows: *Mean* ($\rho = 0.61$), *Sd* ($\rho = 0.74$), *Cov* ($\rho = 0.66$), *MFD* ($\rho = -0.85$), *MFA* ($\rho = -0.60$), *Fat* ($\rho = 0.69$), *CFA* ($\rho = 0.48$), and *Opaque* ($\rho = 0.34$) (**Figure 4**). No statistically significant correlations were observed between *ABx* and *IntN*, *NFA*, or *RFA*.

Given the observed non-linear correlation particularly between *MFD* and *ABx*, it was concluded that standard linear modeling did not sufficiently capture the complexity of these relationships. All the parameters initially demonstrated significant variability and skewness, particularly *MFD*, which exhibited clearly non-linear distributions across the *ABx* range. Following data preprocessing, the log-transformed parameters as described in the Materials and Methods section demonstrated clearly stabilized variances and markedly reduced skewness, resulting in notably improved linear relationships. This enhancement was particularly evident in the correlation between *ABx* and log-transformed *MFD*, where previously pronounced non-linearities were significantly alleviated, thus confirming the necessity of the log-transformation.

Following these log-transformations, Pearson's product-moment correlation coefficients confirmed strong linear associations between *ABx* and key parameters (e.g., *MFD*: $r = -0.86$). Consequently, parameters verified by these statistical assessments formed the basis for the subsequent multiple regression modeling, and also guided the selection of candidate variables employed in the stepwise regression approach as described in the next paragraph.

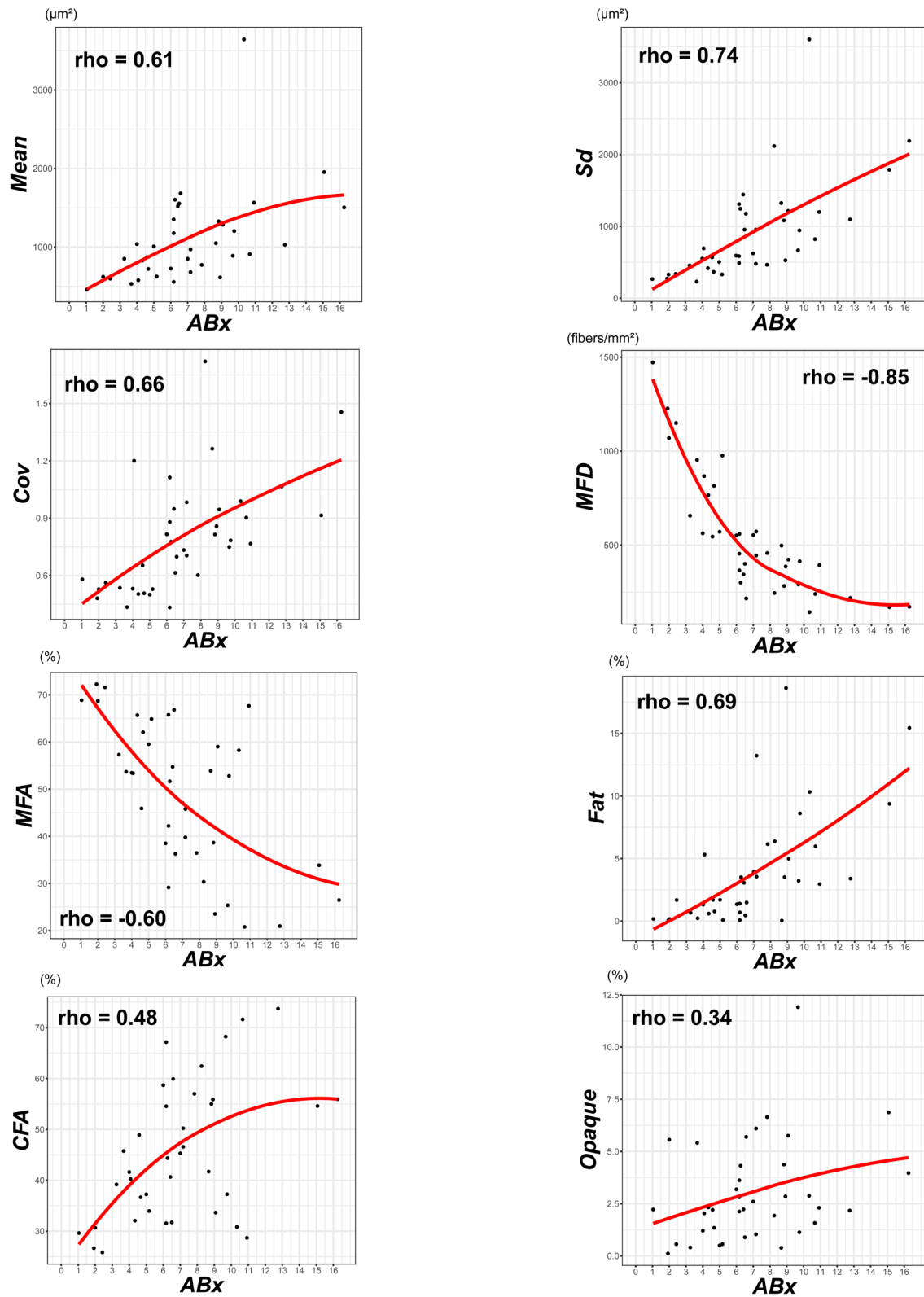


Figure 4: Correlation analysis

Scatter plots demonstrate the correlations between ABx and each parameter, each showing a statistically significant Spearman's correlation coefficient (ρ). Red line is

an approximate curve (R setting: span = 2, except for *MFD* (span = 1)).

(**Abbreviations:** *ABx*, age at biopsy; *Mean*, mean myofiber size; *Sd*, standard deviation of myofiber size; *Cov*, coefficient of variation of myofiber size; *MFD*, myofiber density; *MFA*, myofiber area; *CFA*, connective/fibrotic tissue area; *Fat*, fatty degeneration area; *Opaque*, percentage of opaque fibers relative to *MFD*.)

3.3 Exploratory Analytic Approach

3.3.1 Multiple Regression Analysis and Predictive Error

To explore whether histopathological measures could serve as a tool to assess disease progression in early DMD via age prediction, we selected several variables (log-transformed as described earlier) for multiple regression based on the results of the correlation analyses: *Mean*, *Cov*, *Sd*, *MFD*, *MFA*, *CFA*, and *Fat*. Subsequent multiple regression analyses with stepwise selection excluded *Fat* due to a lack of significant predictive contributions. As a result, two robust regression models were established for predicting *ABx*. The first model, including *MFD* and its interaction term with *Sd*, yielded an adjusted R^2 of 0.79 (**Figure 5**).

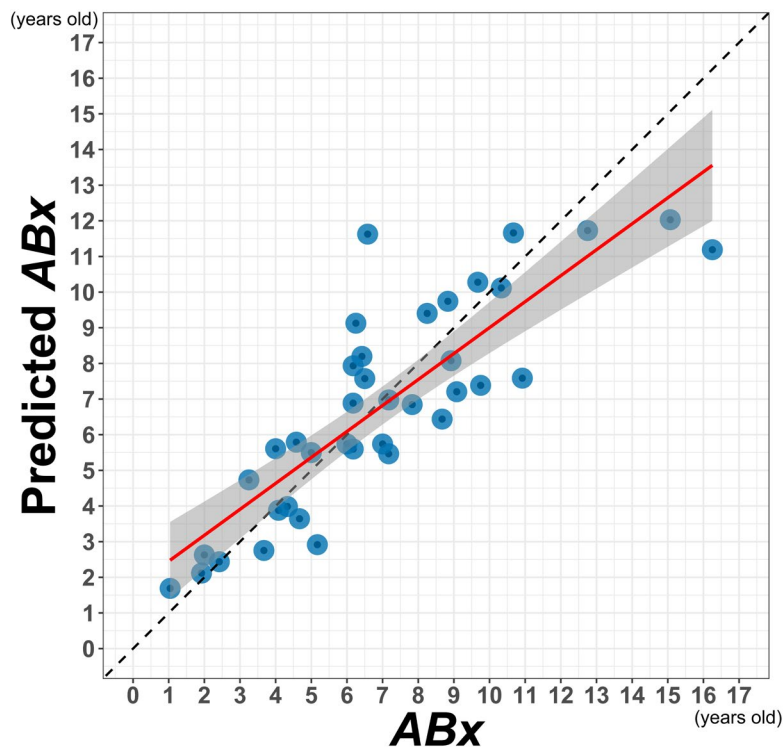


Figure 5: Age-prediction model established by multiple regression analysis
An age-prediction plot was constructed using a multiple regression model with *ABx* as

the dependent variable and *MFD* and its interaction term with *Sd* ($MFD \times Sd$) as predictors. As described in the main text, this model explained approximately 80% of the total variance in the predicted *ABx*. The gray area indicates the standard error. (**Abbreviations:** *ABx*, age at biopsy; *MFD*, myofiber density; *Sd*, standard deviation of myofiber size.)

In this model, the standardized coefficient for *MFD* was -0.87 [CI -1.03 , -0.72] and the interaction term between standardized *MFD* and *Sd* was 0.21 [CI 0.080 , 0.35]; both terms were statistically significant. The second model, including *MFD* and its interaction term with *Mean*, yielded an adjusted R^2 of 0.80 . The general variance inflation factors confirmed the absence of multicollinearity among explanatory variables. Additionally, the significance of the interaction terms was verified through incremental *F*-tests comparing models with and without the interaction terms ($p < 0.01$). The AIC was 53.84 . Diagnostic plots indicated that the first model had superior residual normality and fewer influential outliers (all Cook's distances < 0.50 ; **Figure 6**). The LOOCV for this model yielded an RMSE of 0.49 , with an SD of 0.32 .

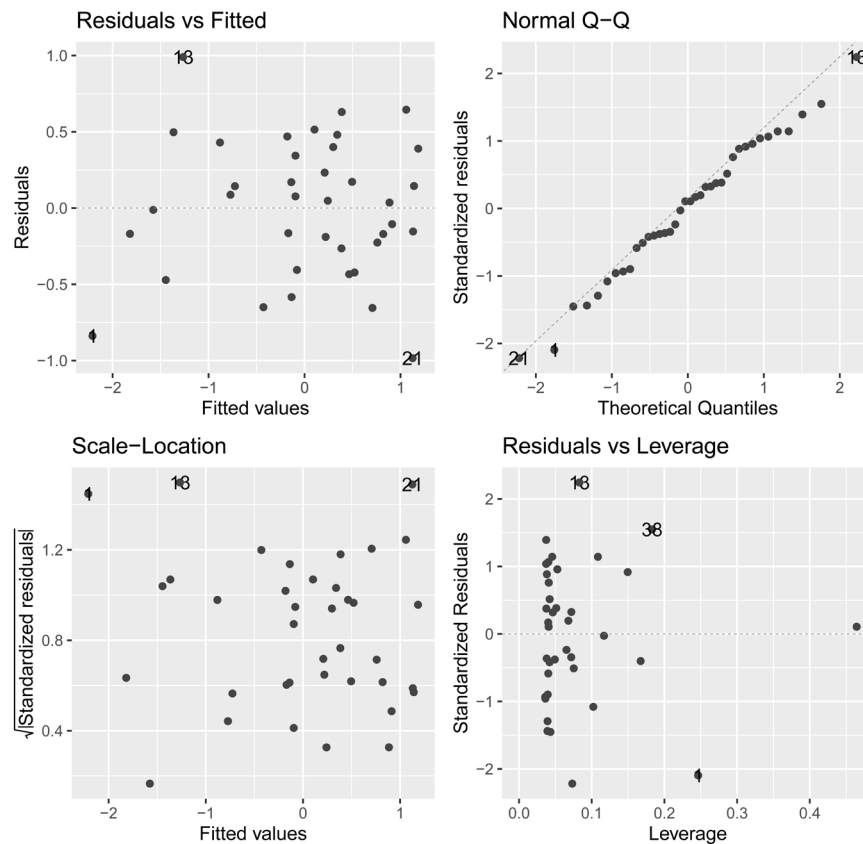


Figure 6: Diagnostic plots for the first multiple regression model

Model diagnostics to evaluate the validity of regression assumptions are shown. The multiple regression model included *ABx* as the dependent variable, and *MFD* and its interaction with *Sd* as predictors.

(**Abbreviations:** *ABx*, age at biopsy; *MFD*, myofiber density; *Sd*, standard deviation of myofiber size.)

Visual inspection of the prediction errors showed age-dependent heterogeneity, with a noticeable deviation around the 6–7-year interval. We therefore compared the distributions of prediction errors across prespecified age groups using a Kruskal–Wallis test (**Figure 7**), noting that this interval aligns with a previously reported morphological transition around ages 6–7 (Peverelli et al., 2015).

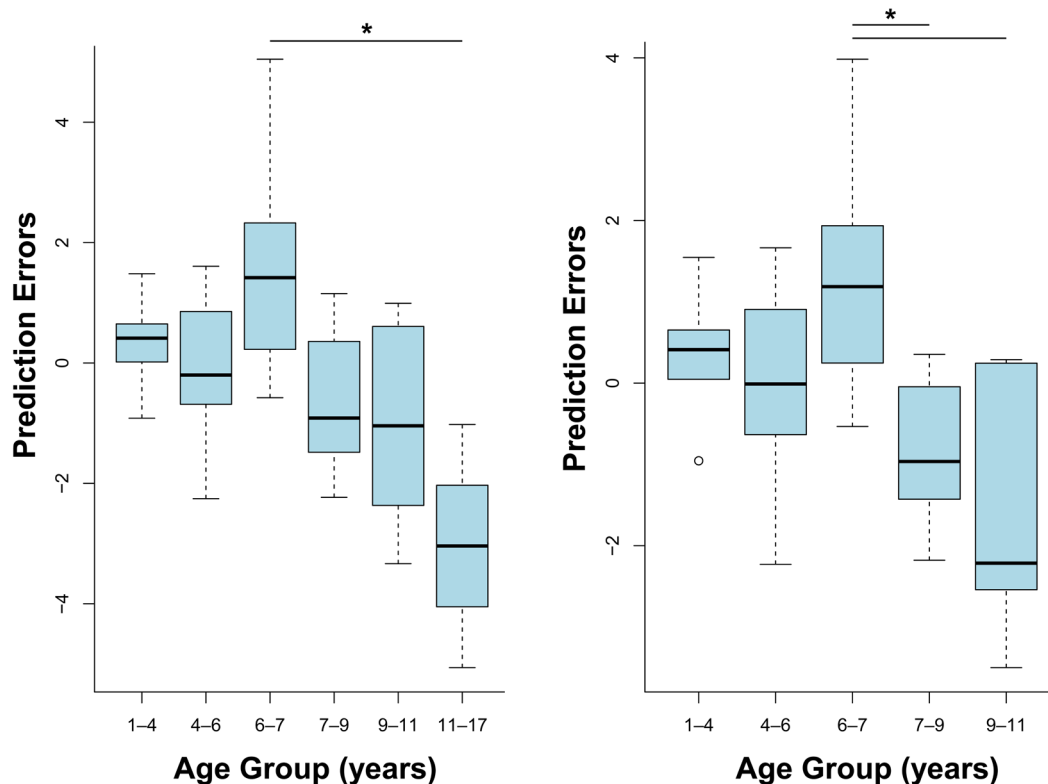


Figure 7: Box plot of residual (predictive error) values

Using the residuals between predicted and actual *ABx* as a variable, an inter-group comparison was performed using the Kruskal–Wallis test (note the limited sample size ($n = 3$) in the 11–17 years age group). The post-hoc test revealed statistically significant differences among age groups ($p < 0.05$; asterisks (*) denote statistical significance).

(**Abbreviation:** *ABx*, age at biopsy.)

Specifically, error analysis showed that the prediction errors were minimized when patients were limited to those younger than 7 years old, whereas prediction accuracy significantly declined in older age groups (**Figure 8**).

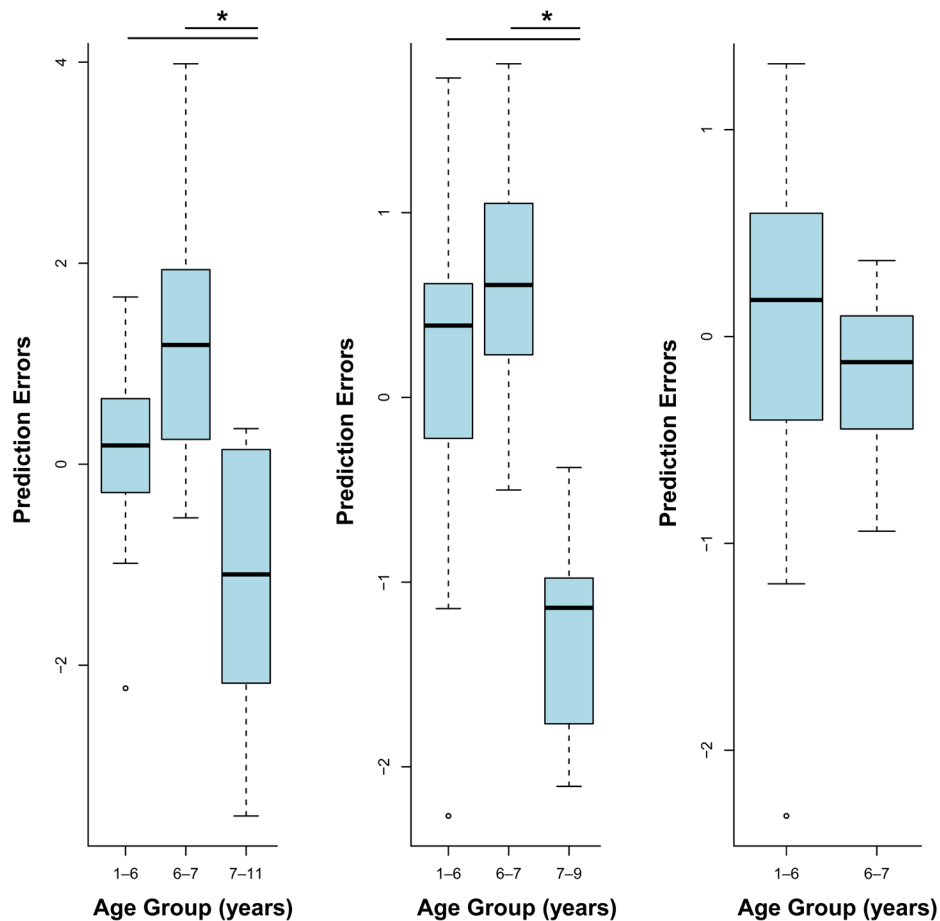


Figure 8: Distribution of prediction errors across different age groups

Boxplots demonstrate prediction errors (the differences between predicted and actual *ABx*) stratified by re-specified age intervals. Statistical comparisons were performed using the Kruskal–Wallis test for three-group analyses (left and center) and Wilcoxon rank-sum test for two-group analysis (right). The post-hoc Conover tests showed significant differences across groups 1–6, 6–7, and 7–11 years (left panel; $p < 0.01$), and similarly significant results were found for age groups 1–6, 6–7, and 7–9 years (center panel; $p < 0.01$). Asterisks (*) denote statistical significances. Conversely, comparison between age groups 1–6 and 6–7 years (right panel) showed no statistically significant difference ($p = 0.30$), with a small effect size as indicated by Cliff's δ of 0.288 [CI $-0.24, 0.69$].

(**Abbreviations:** *ABx*, age at biopsy; CI, Confidence interval.)

Accordingly, in a multiple regression model restricting the sample to patients under 7 years of age (including *MFD* and its interaction with *Sd*; **Figure 9**), both predictors remained statistically significant, showing an improved model fit (adjusted $R^2 = 0.75$, AIC = 29.12) and a comparable predictive accuracy (RMSE = 0.47, SD = 0.38) validated by LOOCV.

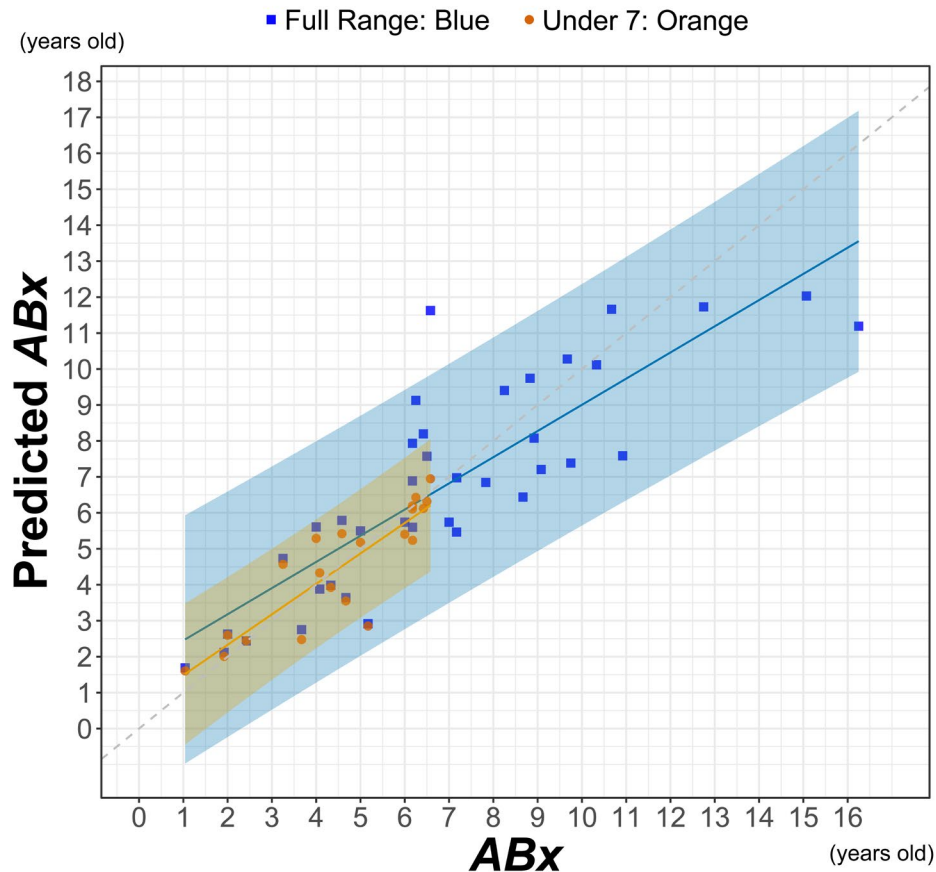


Figure 9: Comparison of predicted *ABx* from two multiple regression models

Scatter plot comparing predicted versus actual *ABx* from two separate multiple regression analyses: a model based on the full dataset (38 samples, indicated by blue squares and blue regression line with CI shaded blue), and a model restricted to patients younger than 7 years (21 samples, indicated by orange dots and orange regression line with CI shaded orange). The narrower CI of the predicted *ABx* values in the restricted dataset corresponds to the reduced prediction error per individual case compared with the model using the full age range. The dotted line represents perfect prediction accuracy ($y = x$).

(Abbreviations: *ABx*, age at biopsy; CI, confidence interval.)

3.3.2 Age-Related Variability in Histopathological Parameters

To further delineate age-related histopathological changes, patients were stratified into age cohorts with a focus on the 6–7-year interval, guided by a previous report (Peverelli et al., 2015) and by our regression and error analyses described above: 1–6 years ($n = 13$), 6–7 years ($n = 8$), and 7–11 years ($n = 14$). Statistically significant differences in the *Mean*, *Sd*, *Cov*, *MFD*, and *CFA* were observed between the 1–6 and 6–7 years age groups (**Figure 10A**, **Table 2**). Comparisons between the 1–6 and 7–11 years age groups showed statistically significant differences across most parameters (**Figure 10A**, **Table 3**), with the exception of *IntN*, *NFA*, *RFA*, and *Opaque*. Notably, box plot analysis showed that *MFD* exhibited the greatest variability within the 1–6 years age group, whereas the *Cov* had the narrowest interquartile-range (0.50–0.56) within this same age group (**Figure 10A**, **Table 3**).

Based on the above results, additional effect sizes were calculated using Hedges' *d* with bootstrapped CI for *Mean*, *Sd*, and *Cov* after excluding extreme outliers affecting normality assumptions (*Mean* and *Sd*: age group, 7–11 years; *Cov*: age groups, 1–6 and 7–11 years). After exclusion, normality assumptions were met for these parameters in the specified age groups, yielding the same statistical significance values as the original findings (**Table 4**).

3.3.3 Age-Stratified MFD Distributions

Although exploratory due to the relatively small sample sizes per group, additional box plots were generated to further elucidate age-related variations in *MFD*, categorizing patients into the following age groups: 1–4 years ($n = 6$), 4–6 years ($n = 7$), 6–7 years ($n = 8$), 7–9 years ($n = 8$), and 9–11 years ($n = 6$) (**Figure 10B**). Notably, the minimum values and first quartiles within the 4–6 years age group overlapped with the maximum values observed in the 6–7 and 7–9 years age groups (**Table 5**).

Among all parameters, *MFD* exceptionally demonstrated near-complete separation in the box plot analyses. Additionally, further comparisons were conducted between the aggregated age groups: the first comparison contrasts *MFD* distributions between the age groups 1–6 ($n = 13$) and 6–11 ($n = 22$) years, whereas the second comparison contrasts age groups 1–7 ($n = 21$) and 7–11 ($n = 14$) years (**Figure 10C**). These plots visually delineated an *MFD* threshold range of approximately 500–600, which subsequently served as the foundation for logistic regression modeling.

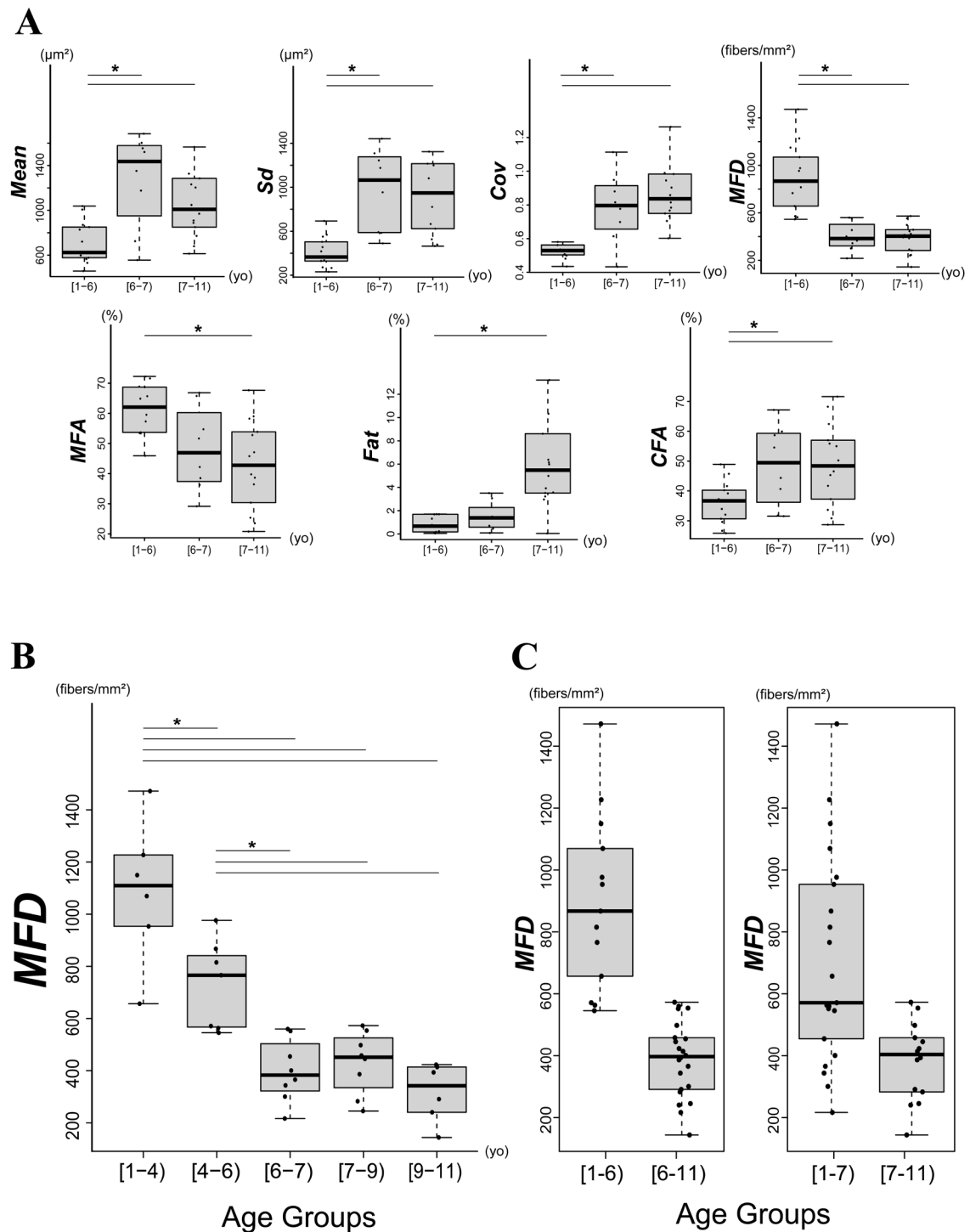


Figure 10: Box plot analyses across age groups

A: Box plots illustrate the results of variance analyses and corresponding post-hoc comparisons for multiple morphological parameters: *Mean*, *Sd*, *Cov*, *MFD*, *MFA*, *Fat*, and *CFA*. Data are stratified by three specific age groups: 1–6, 6–7, and 7–11 years. Each parameter is displayed in a separate panel to facilitate direct visual comparison

among age groups. The age groups are denoted using the interval notation $[X, Y)$, indicating that each range includes ages greater than or equal to X and strictly less than Y . Statistical comparisons between groups were performed using ANOVA, Welch's ANOVA, and Kruskal–Wallis test. Following these analyses, post-hoc Tukey's HSD test, Games–Howell test, or Conover test with Bonferroni correction were applied for pairwise comparisons to determine which specific age groups significantly differed from each other. Statistically significant differences identified by these pairwise comparisons are indicated by asterisks (*). Outliers, defined as data points exceeding the $1.5 \times$ interquartile-range from the quartiles, are not displayed in the plots to maintain visual clarity but were included in all statistical analyses and calculations. These parameters show a quantitative characterization across the age spectrum. In patients within the youngest age group (1–6 years), the variability in parameters related to MFS (*Mean* and *Sd*) as well as MFS heterogeneity (*Cov*) is limited. In contrast, *MFD* values show a wider distribution. Additionally, because the multiple comparisons are exploratory and aimed at characterizing the dynamics of each variable, no multiplicity adjustment was applied at this stage.

B: A refined box plot of *MFD* presents a more granular age stratification (1–4, 4–6, 6–7, 7–9, and 9–11 years), where asterisks (*) denote statistically significant differences between age groups, consistent with panel A. Statistical significance was assessed by first verifying normality and homogeneity of variances using the Shapiro–Wilk and Levene's tests, respectively. ANOVA, followed by the post-hoc Tukey's HSD test, was then conducted to compare groups. For each pairwise comparison, the mean difference and its 95% CI were calculated to determine statistical significance.

C: This panel presents box plots of *MFD* data using two distinct age stratifications for comparative visualization. Specifically, two separate comparisons are depicted: the first comparison contrasts *MFD* distributions between the 1–6-year and 6–11-year age groups (near-complete separation), and the second comparison contrasts the 1–7-year and 7–11-year age groups (obvious overlap). These box plot analyses show two distinct distribution patterns of *MFD* data associated with different age thresholds.

(Abbreviations: MFS, myofiber size; *Mean*, mean MFS; *Sd*, standard deviation of MFS; *Cov*, coefficient of variation of MFS; *MFD*, myofiber density; *MFA*, myofiber area; *CFA*, connective/fibrotic tissue area; *Fat*, fatty degeneration area; ANOVA, analysis of variance; CI, confidence interval.)

Table 3: Variance analysis and post-hoc test results

	Comparison	Effect size [95% CI]	Mean difference [95% CI]	Required sample [§]
<i>Mean</i> *	1–6 vs 6–7 years	0.69 [0.19, 1.00]	556.43 [225.90, 886.96]	8.62
	1–6 vs 7–11 years	0.68 [0.35, 0.92]	498.73 [72.66, 924.80]	8.96
	6–7 vs 7–11 years	−0.29 [−0.80, 0.29]	−57.70 [−578.73, 463.33]	59.53
<i>Sd</i> *	1–6 vs 6–7 years	0.88 [0.65, 1.00]	564.96 [276.94, 852.98]	4.37
	1–6 vs 7–11 years	0.86 [0.63, 0.99]	735.35 [270.39, 1200.30]	4.83
	6–7 vs 7–11 years	−0.054 [−0.57, 0.45]	170.38 [−369.44, 710.21]	1739.43
<i>Cov</i> *	1–6 vs 6–7 years	0.60 [0.039, 1.00]	0.20 [0.013, 0.40]	12.3
	1–6 vs 7–11 years	0.86 [0.56, 1.00]	0.34 [0.14, 0.53]	4.83
	6–7 vs 7–11 years	0.21 [−0.32, 0.68]	0.13 [−0.090, 0.35]	107.15
<i>MFD</i> †	1–6 vs 6–7 years	−1.75 [−2.85, −1.25]	−495.60 [−682.45, −308.75]	6.27
	1–6 vs 7–11 years	−1.81 [−2.91, −1.35]	−512.81 [−689.04, −336.58]	5.94
	6–7 vs 7–11 years	−0.14 [−1.24, 0.92]	−17.21 [−129.64, 95.21]	753.79
<i>MFA</i> ‡	1–6 vs 6–7 years	−1.19 [−2.65, −0.33]	−13.20 [−26.86, 0.48]	12.13
	1–6 vs 7–11 years	−1.51 [−2.56, −0.86]	−18.52 [−30.24, −6.80]	8.01
	6–7 vs 7–11 years	−0.36 [−1.28, 0.46]	−5.33 [−18.82, 8.16]	123.40
<i>Fat</i> *	1–6 vs 6–7 years	0.25 [−0.23, 0.71]	0.39 [−0.82, 1.61]	78.28
	1–6 vs 7–11 years	0.79 [0.44, 1.00]	5.42 [2.64, 8.20]	6.10
	6–7 vs 7–11 years	0.80 [0.46, 1.00]	5.03 [2.19, 7.87]	5.84
<i>CFA</i> †	1–6 vs 6–7 years	1.77 [0.44, 3.87]	12.54 [1.78, 23.31]	6.14
	1–6 vs 7–11 years	1.81 [0.77, 3.71]	12.85 [4.35, 21.35]	5.91
	6–7 vs 7–11 years	0.023 [−1.08, 1.06]	0.31 [−12.17, 12.78]	29960.25

(**Abbreviations:** CI, confidence interval; *Mean*, mean myofiber size; *Sd*, standard deviation of myofiber size; *Cov*, coefficient of variation of myofiber size; *MFD*, myofiber density; *MFA*, myofiber area; *Fat*, fatty degeneration area; *CFA*, connective/fibrotic tissue area.)

*Cliff's δ for variables violating normality assumptions (*Mean*, *Sd*, *Cov*, and *Fat*).

† Glass's Δ for variables with unequal variances (*MFD* and *CFA*).

‡ Hedges' d for normally distributed variable with equal variances (*MFA*).

§ The per-group sample size required to detect the observed effect with 80% power.

Table 4: Additional effect size analyses excluding extreme outliers

	Comparison	Effect size [95% CI]
<i>Mean</i> [†]	1–6 vs 6–7 years	1.81 [0.79, 4.70]
	1–6 vs 7–11* years	1.27 [0.61, 2.26]
	6–7 vs 7–11* years	−0.69 [−2.18, 0.22]
<i>Sd</i> [†]	1–6 vs 6–7 years	2.15 [1.32, 4.43]
	1–6 vs 7–11* years	1.57 [1.10, 2.70]
	6–7 vs 7–11* years	−0.042 [−1.08, 0.71]
<i>Cov</i> [‡]	1–6* vs 6–7 years	1.79 [0.81, 4.24]
	1–6* vs 7–11* years	2.47 [1.87, 4.28]
	6–7 vs 7–11* years	0.36 [−0.60, 1.31]

(**Abbreviations:** CI, confidence interval; *Mean*, mean myofiber size; *Sd*, standard deviation of myofiber size; *Cov*, coefficient of variation of myofiber size.)

*Age groups including extreme outliers.

[†] For *Mean* and *Sd*, data from subject 33 were excluded.

[‡] For *Cov*, data from subjects 8 and 26 were excluded. Although the original effect size analysis for *Cov* between the 1–6 and 6–7 age groups had a CI crossing zero, this was resolved upon exclusion of the specified outliers.

Table 5: Quartile summary of *MFD* across the detailed age groups

Age group	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1–4 years	656.84	982.48	1109.62	1088.09	1207.73	1471.93
4–6 years	545.49*	567.15*	765.96	729.22	841.22	976.36
6–7 years	216.62	332.99	383.21	399.26	479.00	559.89*
7–9 years	245.37	360.62	451.66	430.31	511.88	572.69*
9–11 years	143.94	253.14	342.28	317.69	408.97	423.02

(**Abbreviation:** *MFD*, myofiber density.)

*The minimum value and first quartile within the 4–6-year age group overlapped with the maximum values observed in the 6–7 and 7–9 groups.

3.3.4 MFD Thresholds in Logistic Regression Analysis

Given that the preceding analyses consistently showed age 6 to be the presumed critical threshold, logistic regression was performed to evaluate thresholds within the 500–600 range. Several cutoffs were significant, with low AIC values, high accuracy, and low ORs (**Table 6**). Specifically, *MFD* thresholds between 546 and 551 demonstrated statistical significance (AIC = 29.83, accuracy = 85.71%, OR = 0.019 [CI 0.00087, 0.13]), as did thresholds of 552–553 (AIC = 26.95, accuracy = 88.57%, OR = 0.013 [CI 0.00059, 0.099]) and 554–559 (AIC = 25.27, accuracy = 91.43%, OR = 0.0083 [CI 0.00034, 0.070]). The most robust findings were observed for thresholds between 560 and 563, which yielded the lowest AIC value (19.19), highest classification accuracy (94.29%), and an exceptionally low OR (0.0040 [CI 0.00010, 0.043]). Thresholds of 564–571 (AIC = 24.47, accuracy = 91.43%, OR = 0.0087 [CI 0.00035, 0.074]) and the discrete threshold of 572 (AIC = 28.79, accuracy = 88.57%, OR = 0.014 [CI 0.00064, 0.11]) retained statistical significance.

However, the *p*-value for the *MFD* threshold of 546–551 fibers/mm² slightly exceeded the Bonferroni-corrected significance level of 0.000495, and *MFD* thresholds within the 500–545 and 573–600 ranges failed to achieve statistical significance (**Table 6**). Additional analyses incorporating alternative critical ages (4, 5, or 7 years) identified statistically significant *MFD* thresholds; however, these thresholds encompassed broader ranges than those identified at age 6. The reference age used in this logistic regression analysis was subsequently used, together with the moving-average plot, to set the initial breakpoint for the next segmented regression analysis performed next.

Table 6: Logistic regression analysis results for patients <6 years and ≥6 years

<i>MFD</i> threshold (fibers/mm ²)	AIC	Accuracy	OR [95% CI]
500–545	22.55	88.57%	NS
546–551* (<i>p</i> = 0.00071)	29.83	85.71%	0.019 [0.00087, 0.13]
552, 553 (<i>p</i> < 0.000495)	26.95	88.57%	0.013 [0.00059, 0.099]
554–559 (<i>p</i> < 0.000495)	25.27	91.43%	0.0083 [0.00034, 0.070]
560–563 (<i>p</i> < 0.000495)	19.19	94.29%	0.0040 [0.00010, 0.043]
564–571 (<i>p</i> < 0.000495)	24.47	91.43%	0.0087 [0.00035, 0.074]
572 (<i>p</i> < 0.000495)	28.79	88.57%	0.014 [0.00064, 0.11]
573–600	22.35	91.43%	NS

(**Abbreviations:** AIC, Akaike's information criterion; OR, odds ratio; CI, confidence interval; *MFD*, myofiber density; NS, not significant.)

**MFD* thresholds 546–551 indicate results with unadjusted *p*-values below 0.05 but above the Bonferroni-corrected significance threshold ($0.05/101 \approx 0.000495$).

3.4 Confirmatory Statistical Analysis

3.4.1 Segmented Regression Analysis

Although the preceding exploratory analyses consistently showed a transition around age 6, the variance analyses and logistic regression analyses pragmatically fixed age 6 (or the 6–7-year interval) as an anchor for defining age bands and selecting candidate thresholds, an interval also consistent with the 6–7-year window of rapid *CFA* increase reported by Peverelli et al. (Peverelli et al., 2015). However, the critical age is a quantity that should be derived statistically from the data, rather than assumed as an analytical premise.

Therefore, to treat the critical age itself as an estimand rather than a pre-fixed reference, we applied segmented regression to statistically derive the breakpoint in the age–*MFD* relationship. In the segmented regression analysis for all parameters, the initial breakpoint was visually estimated as 6 years of age (**Figure 11**). The *MFD* slope before and after this breakpoint was markedly different, demonstrating a rapid decline up to approximately 6 years of age with statistical significance (slope: -170.68 [CI $-213.41, -127.95$]), after which the rate of decline markedly decelerated without statistical significance (slope: -28.31 [CI $-76.47, 19.85$]) (**Figure 12**). The breakpoint itself was precisely estimated at 6.25 years [CI 5.08, 7.42] (**Figure 12**). This analysis yielded a high coefficient of determination (adjusted $R^2 = 0.80$). ANOVA-based comparison of the simple linear and segmented regression models showed that segmented regression provided a statistically improved fit to the data ($F(2, 31) = 10.37$, $p < 0.001$). LOOCV yielded an RMSE of 151.44 compared with the original RMSE of 132.90 (ratio: 1.14), a mean breakpoint of 6.26 years (median: 6.25 years, range: 5.50 to 6.40 years), and an average CI ranging from 5.08 to 7.48 years. Furthermore, when the initial breakpoint seed was varied systematically between 4 and 8 years in increments of 0.5 years, the segmented regression consistently converged on the identical breakpoint estimate of 6.25 years, with a maximum observed deviation of less than 0.003 years. By contrast, other parameters demonstrated no statistical significance in the segmented modeling. Therefore, all subsequent statistical analyses were applied to *MFD*.

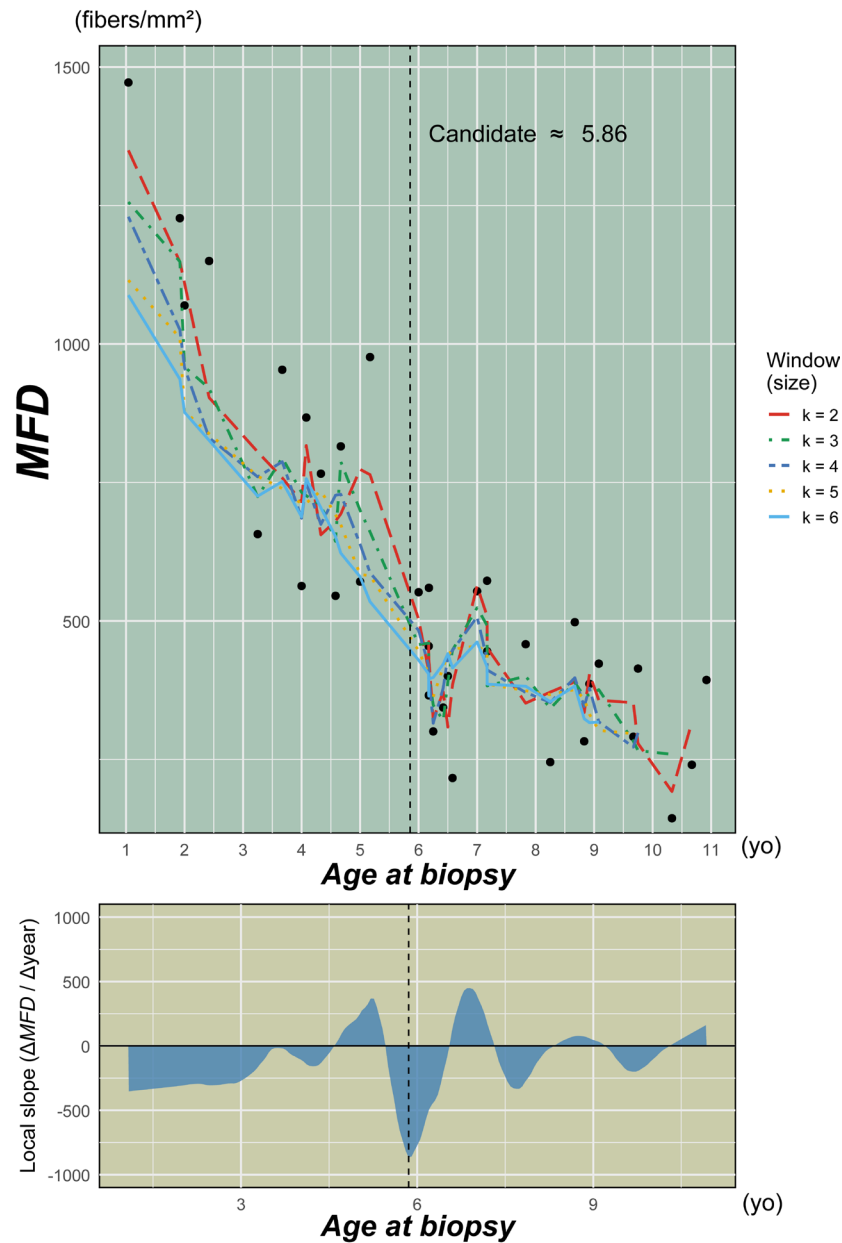


Figure 11: Mean-average plot of *MFD* in relation to age at biopsy

The upper panel shows the relationship between *MFD* and biopsy age, represented by a moving-average curve (window size ranging from $k = 2$ to 6). The vertical dashed line marks the steepest decline in the moving-average plot, showing a candidate breakpoint at 5.86 years. Individual data points correspond to actual patient measurements. The lower panel shows the local slope calculated directly from the moving-average curve (window size: $k = 2$).

(**Abbreviations:** *MFD*, myofiber density; RMSE, root mean squared error.)

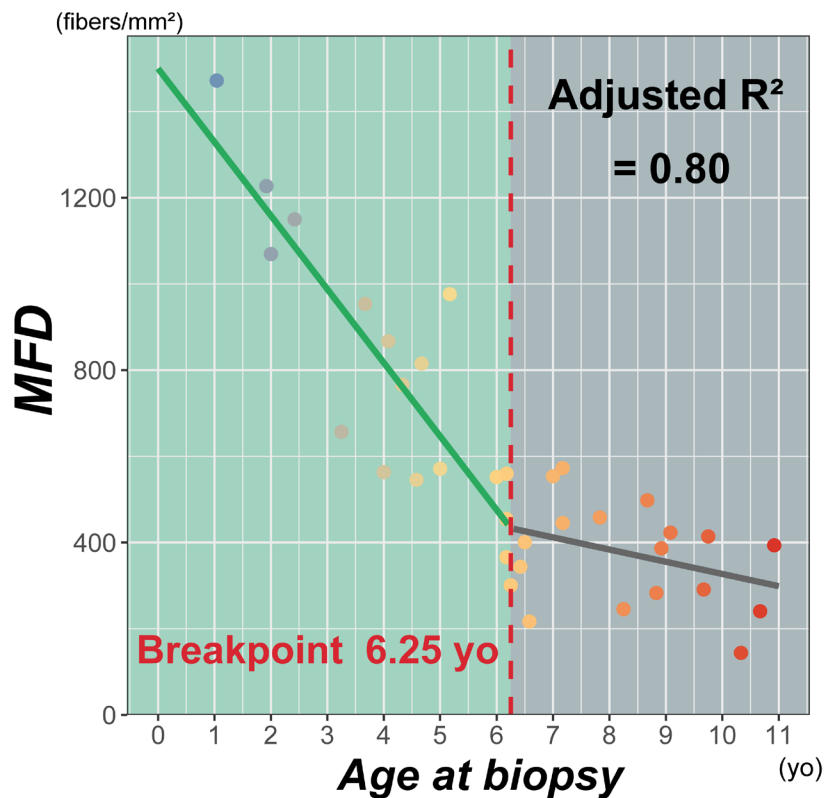


Figure 12: Segmented regression analysis

Segmented regression analysis identified a statistically significant breakpoint at 6.25 years (indicated by the vertical dashed line). Prior to this age, *MFD* exhibited a rapid decline, after which the rate of decline noticeably leveled off.

(**Abbreviation:** *MFD*, myofiber density.)

3.4.2 *MFD Replacement by ATPase-Stained Slide Samples*

The *MFD* values obtained from H&E-, G-T-, or M-T-stained slide samples were exploratorily replaced with those obtained from ATPase-stained slide samples (14 cases). This adjustment also resulted in near-complete separation between age groups before and after 6 years, except for a single case (subject 10) exhibiting a low *MFD* before age 6, with a strong Spearman's rank correlation coefficient for *ABx-MFD* ($\rho = -0.87$) across all 38 cases (**Figure 13**). Additionally, despite this replacement, *MFD* exhibited the same dynamics in the segmented regression analysis, characterized by a rapid decline before age 6 and stability thereafter. Notably, this adjustment identified a similar breakpoint age (6.42 [CI 5.28, 7.56]).

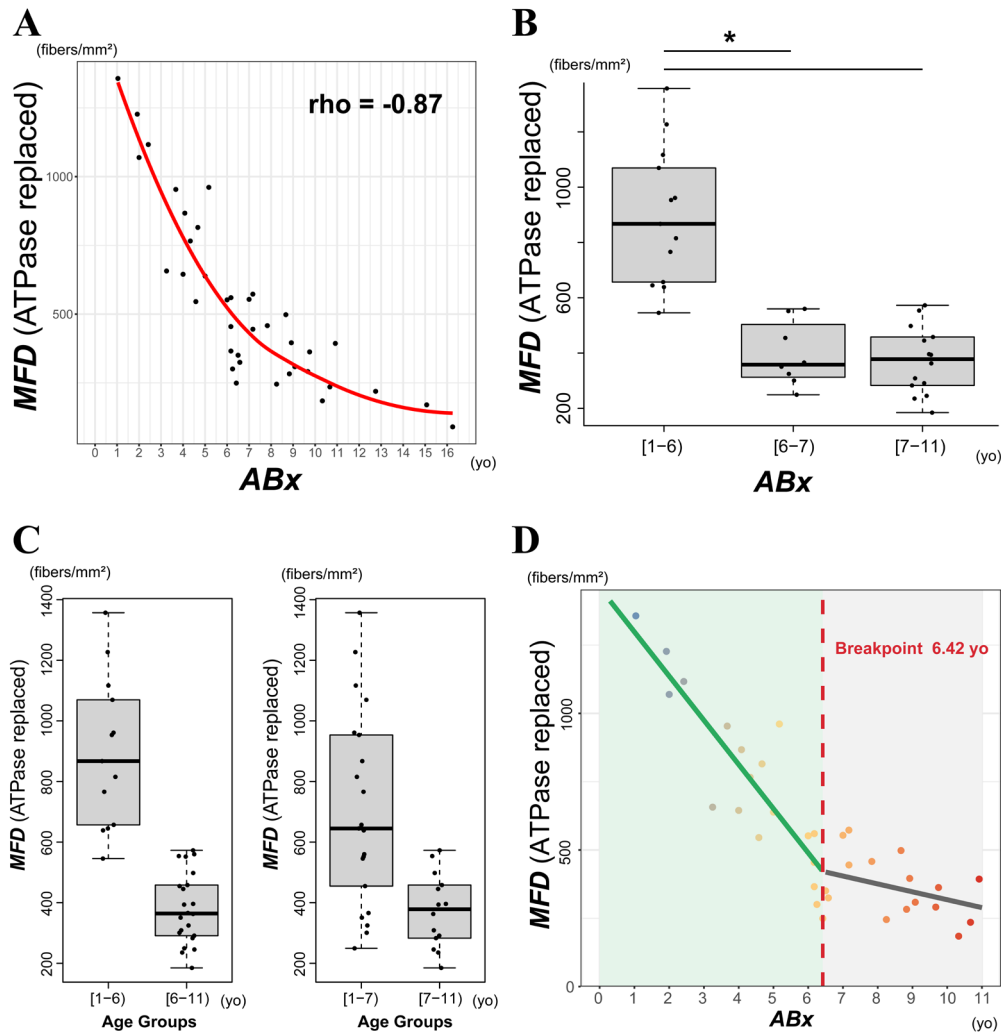


Figure 13: Replacement of *MFD* with ATPase-based measurements

A: Scatter plot demonstrates a strong inverse correlation between *ABx* and *MFD* after replacing original measurements with those derived from ATPase-stained samples.

B: Box plots illustrate ATPase-replaced *MFD* across the age stratification of 1–6, 6–7, and 7–11 years. Asterisk (*) denotes the statistical significances between age groups (for ages 1–6 vs. 6–7 years and 1–6 vs. 7–11 years).

C: Box plots comparing ATPase-replaced *MFD* using two different age-threshold cutoffs are shown: 1–6 vs. 6–11 years and 1–7 vs. 7–11 years. By presenting data in this comparative manner, this panel highlights age 6 as a critical histopathological threshold.

D: Segmented regression analysis demonstrates a statistically significant breakpoint at age 6.42 years (vertical dashed line).

(**Abbreviations:** ATPase, actomyosin adenosine triphosphatase; *ABx*, age at biopsy; *MFD*, myofiber density.)

3.4.3 Bayesian Analyses and Monte Carlo Simulations for MFD

The first Bayesian segmented regression analysis indicated a posterior mean slope of -170.44 MFD/year [95% Credible Interval (CrI): -207.47 , -133.58] before the breakpoint (age, 6.25 years), and that of -28.42 MFD/year [CrI: -65.22 , 8.77] after the breakpoint. Using a ROPE of ± 62.00 MFD/year, derived from the difference between original and ATPase-replaced MFD values, no posterior estimates (0%) fell within this ROPE prior to the breakpoint, whereas nearly all posterior estimates (96.26%) fell within this ROPE after the breakpoint. Bayesian model comparison based on BIC yielded values of approximately 470.20 for the single-line regression model and approximately 455.82 for the segmented regression model, corresponding to an approximate BF_{10} of 1324. The sensitivity analysis using ROPE width defined as $+1SD$, $+1.5SD$, and $+2SD$ yielded posterior probabilities $>96\%$ (post-breakpoint) and $<0.5\%$ (pre-breakpoint), closely consistent with the aforementioned findings.

The Monte Carlo Method (2×10^6 replicates) quantified the probability that an MFD-based rule would place a muscle biopsy on the wrong side of the 6.25-year boundary. When the true age difference was 0.25 years ($\Delta MFD = 42.67$), the mean misclassification probability was 38.92%; this probability decreased monotonically to 1.21% when the age gap reached 2 years ($\Delta MFD = 341.36$), and the error rates for "<6.25 years" and " ≥ 6.25 years" assignments remained symmetric throughout this simulation (Table 7).

Table 7: Monte Carlo simulation (2×10^6 replicates) of mean mis-classification

ΔABx	ΔMFD (fibers/mm ²)	Mis-classification <6 years	Mis-classification ≥ 6 years	Mean Error
0.25 years	42.67	38.96%	38.87%	38.92%
0.5 years	85.34	28.67%	28.65%	28.66%
0.75 years	128.01	19.93%	19.85%	19.89%
1 years	170.68	12.97%	13.01%	12.99%
1.25 years	213.35	7.93%	7.97%	7.95%
1.5 years	256.02	4.56%	4.56%	4.56%
1.75 years	298.69	2.43%	2.42%	2.43%
2 years	341.36	1.21%	1.22%	1.21%

(Abbreviations: ABx, age at biopsy; MFD, myofiber density.)

The second Bayesian posterior analysis, using weakly informative priors for the *MFD* segmented regression model, yielded an estimated breakpoint at 6.37 years [CrI: 5.24, 7.66] (**Table 8**). Posterior slope estimates revealed a significant decline of -165.14 *MFD*/year [CrI: -200.32 , -131.15] in the *MFD* before this breakpoint, whereas the rate of decline substantially decreased to -28.72 *MFD*/year [CrI: -67.84 , 15.94] after the breakpoint, which was not statistically significant (**Table 8**). Bayesian model comparison based on LOOIC indicated that segmented regression provided superior predictive performance to the simpler linear model, with an expected log predictive density difference of 7.31. Additionally, the posterior slope preceding the breakpoint fell completely outside the ROPE interval, whereas the slope following the breakpoint largely overlapped with the ROPE (95.01% of the posterior distribution within ROPE; **Figure 14**).

Table 8: Bayesian segmented regression model with weakly informative priors*

Parameter	Mean	SD	2.5% CrI	Median	97.5% CrI
Intercept (pre-BP)	1459.61	81.67	1292.3	1460.28	1617.48
Slope (pre-BP) [†]	-165.14	17.34	-200.32	-164.87	-131.15
Slope (post-BP) [†]	-28.72	21.09	-67.84	-29.73	15.94
Breakpoint age (years)	6.37	0.58	5.24	6.37	7.66

(**Abbreviations:** BP, breakpoint; CrI, credible interval; ROPE, Region of Practical Equivalence; SD, standard deviation; *ABx*, age at biopsy; *MFD*, myofiber density.)

Weakly informative priors were defined based on estimates from the original segmented regression model as follows: β_0 (Intercept): normal(1500, 300); β_1 (pre-breakpoint slope): normal(-171 , 40); β_Δ (Difference between slopes): normal(142, 40); ψ^ (Breakpoint age; centered at zero): normal(0, 2). The prior SD for the intercept was set to 300, corresponding to the SD of *MFD*. The error term was assumed as $\varepsilon_i \sim N(0, \sigma^2)$.

[†] The proportion of the posterior distribution lying within the ROPE was 0% for the slope before the breakpoint, strongly supporting a non-zero slope, whereas the posterior distribution after the breakpoint largely overlapped with the ROPE (95.01%), indicating practical equivalence to zero.

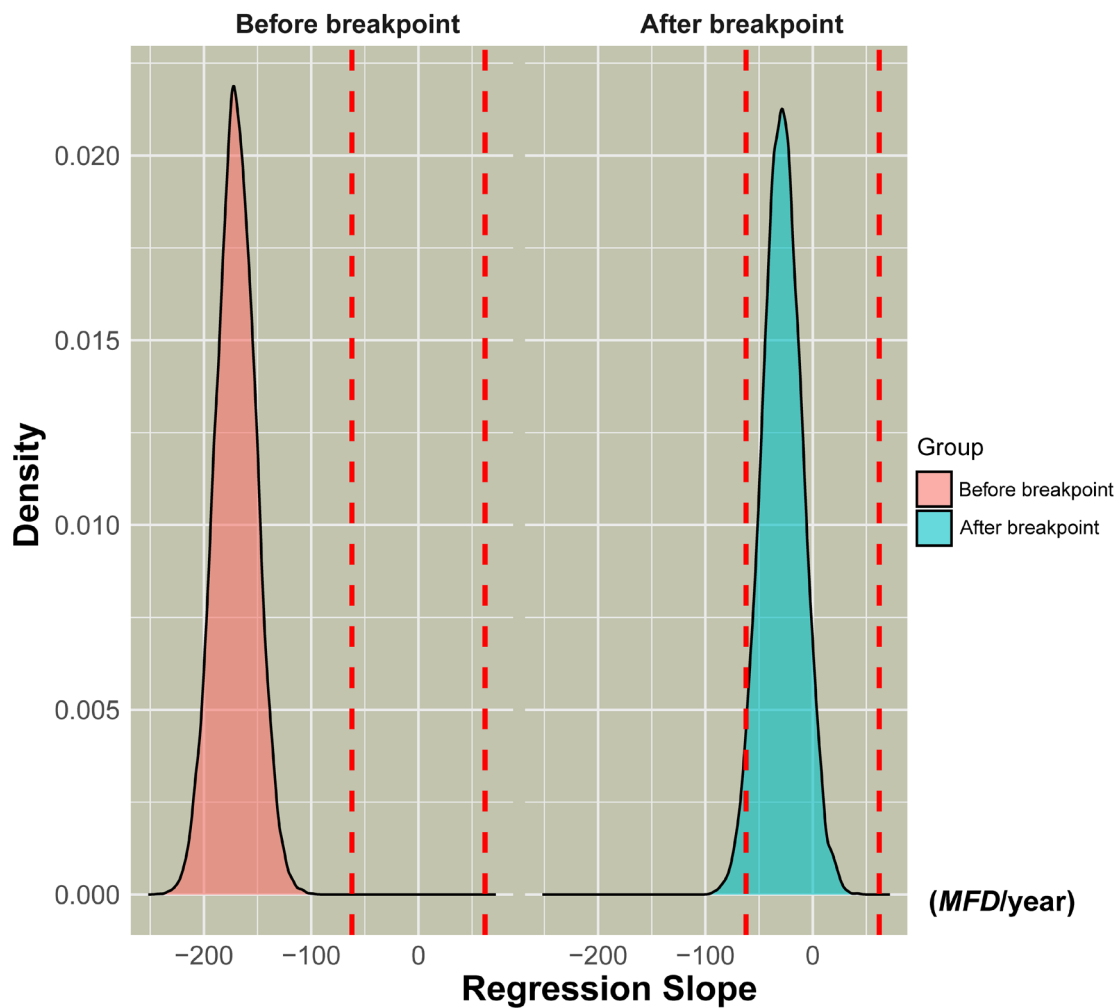


Figure 14: ROPE analysis

This plot shows the posterior distributions of regression slopes derived from the Bayesian segmented regression analysis, separated at the breakpoint. The posterior distribution of slopes before the breakpoint is shown on the left (red), whereas the distribution of slopes after the breakpoint is depicted on the right (blue). The dashed red lines represent the ROPE, defined as ± 62.00 *MFD*/year.

(**Abbreviations:** ROPE, Region of Practical Equivalence; *MFD*, myofiber density.)

Monte Carlo sample-size simulations indicated that the current study's sample size ($n = 35$) provided a probability of 72.43% for detecting the breakpoint within ± 1 year, 79.47% within ± 1.25 years, and 84.54% within ± 1.50 years (**Table 9**). In addition, probabilities corresponding to a range of alternative sample sizes, as determined by the Monte Carlo simulations, are also presented in **Table 9**.

Table 9: The results of sample-size simulations by Monte Carlo method*

Sample size	Precision (± 1 year) [†]	Precision (± 1.25 years) [†]	Precision (± 1.5 years) [†]
20	58.18%	65.43%	72.66%
25	65.65%	71.01%	76.77%
30	69.80%	76.43%	82.08%
35	72.43%	79.47%	84.54%
40	75.11%	81.64%	86.44%
45	77.00%	84.08%	88.16%
50	79.71%	86.14%	89.96%
55	81.59%	87.63%	91.58%
60	83.42%	89.01%	92.40%
65	84.51%	89.81%	93.87%
70	85.61%	90.65%	94.01%

* For each replicate, values of age at biopsy were obtained through bootstrap resampling from the cohort younger than 11 years, while model parameters (β_1 , β_Δ , σ , ψ) were independently sampled from their respective Bayesian posterior distributions ($n = 10,000$ draws). A replicate was considered successful if the breakpoint determined by segmented regression was within ± 1.0 – 1.5 years of the true breakpoint.

[†] Precision columns indicate the Monte Carlo-based probability that the estimated breakpoint was within ± 1.0 – 1.5 years of the posterior-derived true breakpoint.

When the fitted logistic model established in the previous section was tested in terms of the conditional probability of being < 6.25 years (rather than overall classification accuracy), three *MFD* cut-offs emerged. This regression analysis identified thresholds such that values ≥ 595.51 fibers/ mm^2 corresponded to an approximately 80% probability of being younger than 6.25 years, values between 426.36 and 595.51 fibers/ mm^2 defined a grey zone with approximately 50% probability, and values ≤ 426.36 fibers/ mm^2 corresponded to an approximately 80% probability of being 6.25 years or older. Application of these cut-points allocated 10 cases to the high-probability "<6.25 years" zone, 11 to the grey zone, and 14 to the low-probability " ≥ 6.25 years" zone ($ABx < 11$, $n = 35$; **Table 10**).

Table 10: Zone summary and *MFD* threshold

<i>MFD</i> zone	<i>MFD</i> threshold (fibers/mm ²)	Cases	Presumed therapeutic window*
High (≥80%, <6.25 years)	595.51	10	Large therapeutic window
Grey (20–80%, <6.25 years)	426.36–595.51	11	Requires individual judgement
Low (≥80%, ≥6.25 years)	426.36	14	Small therapeutic window

(**Abbreviation:** *MFD*, myofiber density.)

*Of note, the presumed therapeutic windows described here represent hypothetical classifications based solely on *MFD* thresholds derived from logistic regression probabilities; clinical validation of these categories was not performed in this analysis.

3.4.4 Revised Variance Analysis

Group-wise variance analysis showed significant pair-wise differences in *Mean*, *Sd*, and *Fat* between the Early band and every older band, and a significant difference in *MFA* and *Cov* only between the Early and Late bands (**Figure 15**). In the Early–Late comparison, Cliff's δ indicated large effects for *Mean* ($\delta = 0.80$ [CI: 0.49, 1.00]), *Sd* ($\delta = 0.88$ [CI: 0.65, 1.00]), *Cov* ($\delta = 0.83$ [CI: 0.49, 1.00]), and *Fat* ($\delta = 0.75$ [CI: 0.36, 1.00]), while Hedges' d showed a pronounced decrease in *MFA* ($d = -1.36$ [CI: -2.54, -0.65]) and a large increase in *CFA* ($d = 1.04$ [CI: 0.31, 2.16]). However, *CFA* did not attain statistical significance in the variance analysis. The complete statistics, with *MFD* itself included as a positive control, are provided in **Table 11**.

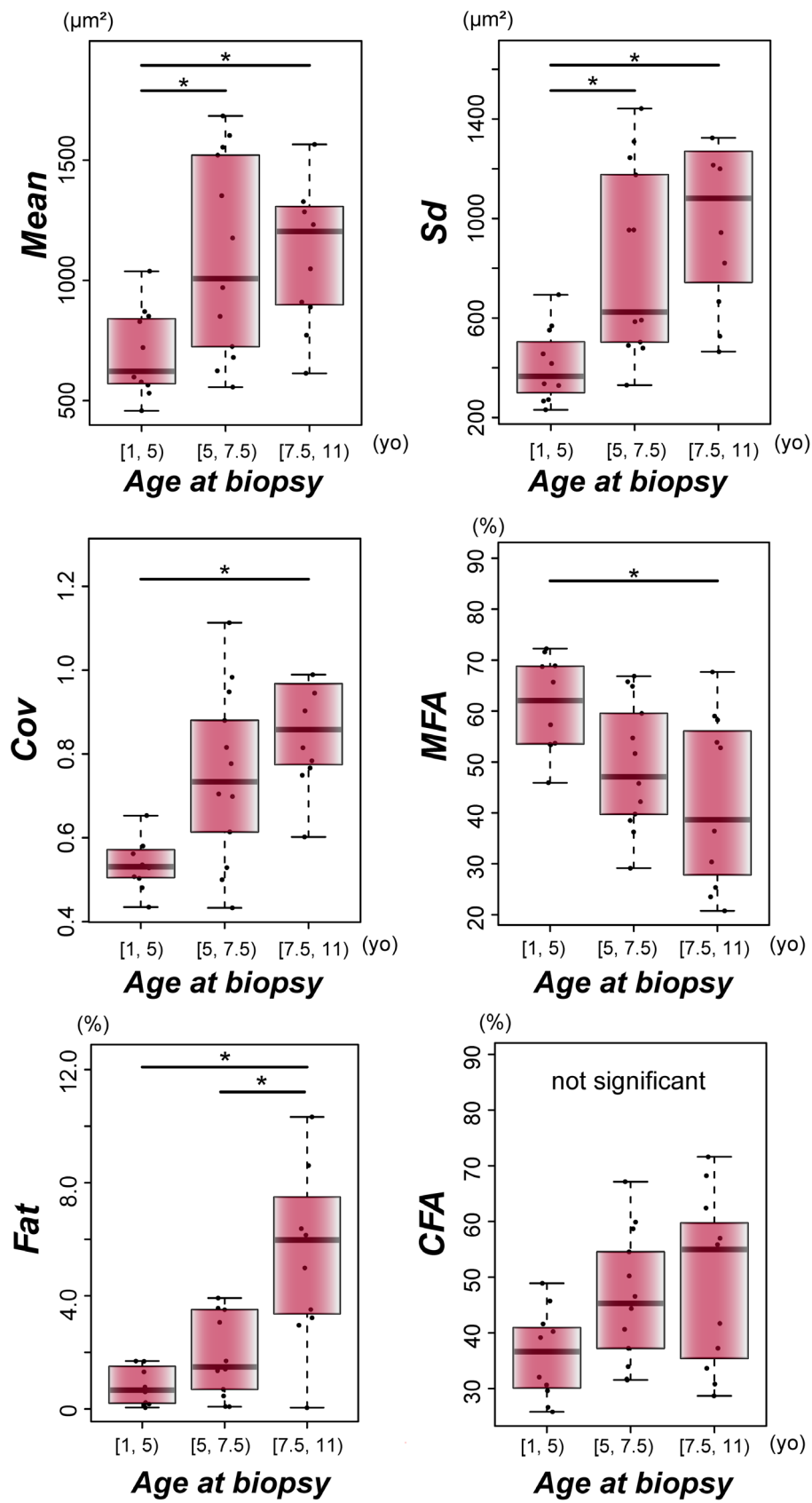


Figure 15: Box plots of revised variance analyses

Box plots summarize revised variance analyses and post-hoc comparisons for morphological parameters (*Mean*, *Sd*, *Cov*, *MFA*, *Fat*, *CFA*) across three age groups: 1–5, 5–7.5, and 7.5–11 years. Brackets [X, Y) denote ages X or older but younger than Y. * indicates statistical significance (Analysis of Variance/Kruskal–Wallis) adjusted by false-discovery rate. Outliers omitted visually but included in calculations. (Abbreviations: *Mean*, mean myofibre size; *Sd*, standard deviation of myofibre size; *Cov*, coefficient of variation of myofibre size; *MFA*, myofibre area; *CFA*, connective/fibrotic tissue area; *Fat*, fatty degeneration area.)

Table 11: Revised variance analysis, effect size, and adjusted *p*-value with power

	Comparison	Effect size [95% CI]	Mean difference [95% CI]	<i>p</i> -value [‡]	Required [§]
<i>Mean</i>*	1–5 vs 5–7.5 years	0.61 [0.22 to 0.92]	403.72 [145.14 to 662.29]	<0.01	11.74
	1–5 vs 7.5–11 years	0.80 [0.49 to 1.00]	620.75 [93.48 to 1148.02]	<0.01	5.88
	5–7.5 vs 7.5–11 years	0.09 [–0.38 to 0.57]	217.03 [–345.70 to 779.77]	0.55	602.95
<i>Sd</i>*	1–5 vs 5–7.5 years	0.73 [0.38 to 0.97]	414.10 [180.25 to 647.95]	<0.01	7.43
	1–5 vs 7.5–11 years	0.88 [0.65 to 1.00]	861.90 [288.90 to 1434.90]	<0.0001	4.38
	5–7.5 vs 7.5–11 years	0.30 [–0.16 to 0.73]	447.80 [–154.31 to 1049.91]	0.18	53.58
<i>Cov</i>*	1–5 vs 5–7.5 years	0.48 [0.007 to 0.89]	0.16 [–0.019 to 0.33]	0.039	19.74
	1–5 vs 7.5–11 years	0.83 [0.49 to 1.00]	0.35 [0.12 to 0.59]	<0.001	5.24
	5–7.5 vs 7.5–11 years	0.40 [–0.06 to 0.79]	0.20 [–0.03 to 0.42]	0.064	29.75
<i>MFD</i>[†]	1–5 vs 5–7.5 years	–1.89 [–3.06 to –1.01]	–431.25 [–641.04 to –221.46]	<0.0001	6.35
	1–5 vs 7.5–11 years	–2.71 [–3.97 to –1.89]	–573.52 [–791.87 to –355.16]	<0.0001	3.75
	5–7.5 vs 7.5–11 years	–0.88 [–1.69 to –0.26]	–142.27 [–352.06 to 67.52]	0.23	21.43
<i>MFA</i>[†]	1–5 vs 5–7.5 years	–1.05 [–2.08 to –0.33]	–11.78 [–24.75 to 1.20]	0.12	15.39
	1–5 vs 7.5–11 years	–1.36 [–2.54 to –0.65]	–18.74 [–32.25 to –5.24]	0.015	9.52
	5–7.5 vs 7.5–11 years	–0.47 [–1.42 to 0.33]	–6.97 [–19.94 to 6.01]	0.4	72.31
<i>Fat</i>*	1–5 vs 5–7.5 years	0.36 [–0.10 to 0.78]	1.50 [–0.69 to 3.70]	0.15	37.6
	1–5 vs 7.5–11 years	0.75 [0.36 to 1.00]	5.28 [2.04 to 8.52]	<0.01	6.98
	5–7.5 vs 7.5–11 years	0.57 [0.12 to 0.94]	3.78 [0.12 to 7.44]	0.026	13.83
<i>CFA</i>[†]	1–5 vs 5–7.5 years	0.99 [0.31 to 1.92]	10.19 [–1.79 to 22.17]	0.16	16.89
	1–5 vs 7.5–11 years	1.04 [0.31 to 2.16]	13.19 [0.72 to 25.66] [¶]	0.11 [¶]	15.52
	5–7.5 vs 7.5–11 years	0.21 [–0.62 to 1.13]	3.00 [–8.98 to 14.98]	0.81	336.47

(**Abbreviations:** *CI*, confidence interval; *CFA*, connective/fibrotic tissue area; *Cov*, coefficient of variation of myofiber size; *Fat*, fatty degeneration area; *Mean*, mean myofiber size; *MFA*, myofiber area; *MFD*, myofiber density; *Sd*, standard deviation of myofiber size.)

* Cliff's δ was used for variables that did not meet normality assumptions.

† Hedges' d was used for variables meeting assumptions of normality and equal variances.

‡ P -values were adjusted using the false-discovery rate method.

§ Required sample size per group to detect the observed effect with 80% statistical power.

¶ The CI excludes zero; however, the adjusted p -value was 0.11 (>0.05), and thus the result was considered not statistically significant.

4. Discussion

4.1 Positioning of the Current Findings and Image-Analytic Innovations

4.1.1 Overview

Two novel insights emerge from the present study: (1) *MFD* serves as a candidate surrogate biomarker for early-stage DMD progression, and (2) the identified age 6 threshold and the observed *MFD* dynamics delineate a critical histopathological transition in disease pathology. These findings suggest a previously unrecognized subclinical phase of accelerated myofiber loss in the early stages of DMD. By capturing early tissue alterations undetectable by conventional indices, these results provide a potential framework for defining the optimal timing of corticosteroid initiation and other emerging therapies, thereby informing more individualized therapeutic decisions.

While previous research predominantly used *MFA* along with connective tissue area or fibrotic changes as indices (Desguerre et al., 2009; Peverelli et al., 2015), this study has revealed that the number of myofibers per square millimeter provides a more dynamic and earlier indicator of pathological changes. This parameter is proposed as muscle fiber "density." Specifically, the current analysis reveals a rapid decline in *MFD* before the age of 6 years, challenging the previous understanding that muscular histological stability is maintained until the age of 6–7 years (Peverelli et al., 2015).

4.1.2 Morphometric Methodologies and Image Analyses

A distinguishing feature of this study is the use of archival muscle biopsy specimens collected over four decades ago. Although the primary analysis of this study focused on H&E-, G-T-, and M-T-stained slides, severely faded ATPase-stained specimens were also found, as exemplified by the specimen shown in **Figure 1A**, highlighting the extent of degradation observed. The application of LUT-based digital restoration techniques underscores the feasibility of recovering highly compromised histological materials at minimal cost. Notably, in an exploratory analysis, *MFD* values were substituted from 14 cases initially assessed via H&E/G-T/M-T staining with those derived from ATPase-stained slides. Despite this replacement, *MFD* exhibited the same dynamics, characterized by a rapid decline before age 6 and stability thereafter. Notably, this adjustment resulted in near-complete separation between age groups before and after 6 years, except for a single case (subject 10) exhibiting a low *MFD* before age 6, with a strong Spearman's rank correlation coefficient for *ABx*–*MFD* ($\rho = -0.87$) across all 38 cases (**Figure 13**). This reinforces both the robustness of *MFD* as a new

metric and the efficacy of this color-restoration approach. Additionally, Focus Stacking was employed to mitigate focal plane variations attributable to uneven sectioning or mounting, facilitating the generation of high-resolution, well-focused images from suboptimal histological preparations. These methodologies hold broader implications for various pathology fields because they enable the partial revitalization of archival specimens compromised by diverse adverse conditions, thereby expanding opportunities for histopathological research and historical dataset utilization.

Historically, MFS quantification in DMD histopathological research has often relied on semi-quantitative methods, while Bell and Conen employed the *Cov* to quantitatively evaluate MFS changes in 1967 under the technological constraints of that era (Bell and Conen, 1967; Desguerre et al., 2009; Peverelli et al., 2015). In this study, because an objective and more precise assessment of MFS necessitates the enumeration of all individual myofibers, *MFD* values have inherently emerged as part of the quantitative analytical framework. Notably, despite comparable *MFA*, a 1.5-fold difference in *MFD* was observed between patients aged 4 and 6 years (765.96 vs. 454.67; **Figure 3A, 3B**), underscoring a substantial histopathological distinction. Microscopically, densely packed myofibers may superficially convey an impression of increased *MFD*. However, the observations in this study suggest that mild hypertrophy may already be present before age 6, contributing to a uniform yet globally enlarged myofiber morphology that could obscure early-stage myofiber depletion upon cursory examination. The stability of MFS-related parameters before age 6, particularly the narrow interquartile range of 0.50–0.56 in *Cov* (**Figure 10A, Table 3**), further supports this interpretation. Thus, these findings provide mechanistic insights into the differences in effect sizes observed between *MFA* and *MFD*.

Before the main statistical analysis, a thorough statistical evaluation using paired specimens stained with H&E and G-T methods was conducted, which demonstrated negligible batch effects and a very strong correlation between the two staining techniques. This approach, wherein the correlation coefficients are employed as an alternative method to indirectly indicate that statistically significant differences are unlikely, as conventional statistical tests inherently cannot conclusively prove the absence of differences, is based on the exploratory approach described by Bell and Conen (Bell and Conen, 1967). The results of Bayesian analysis also confirmed the negligible batch effects between H&E and G-T methods. These findings strongly suggest that methodological variability arising from using multiple staining techniques did not significantly influence the results of the quantitative image analyses. Although

comprehensive validation across all the staining methods used in the present study would have been ideal, the rigorous comparative approach employed was considered sufficient to demonstrate methodological robustness and to validate the integration of data derived from multiple staining modalities. Future studies employing larger sample sizes under more rigorously controlled conditions are warranted to further substantiate the reliability and generalizability of the findings.

4.2 From Exploratory to Confirmatory: Statistical Trajectory and Rationale

4.2.1 Exploratory Statistical Approaches

Although related points are noted in the Results, the following considerations are restated here, as they directly affect interpretation of the results.

In statistical analysis, all parameters, including ABx , were treated as continuous variables. Initially, an age-predicting model was planned in order to establish an objective and reproducible method for evaluating muscular status in DMD, followed by correlation analysis for parameter selection. Subsequently, logarithmic transformation of the selected variables and multiple regression analysis with ABx as the dependent variable were performed. However, dividing the predicted values into more detailed age groups revealed that prediction errors varied significantly across specific intervals. The Kruskal–Wallis test was then conducted to determine the age range within which prediction errors were minimized, revealing robust predictions up to approximately 7 years of age. Indeed, limiting the multiple regression to patients aged <7 years notably improved the model fit, as indicated by a marked decrease in the AIC from 53.84 to 29.12. Furthermore, LOOCV confirmed that the restricted model (<7 years) maintained comparable predictive accuracy (RMSE: 0.47) to the full-range model (RMSE: 0.49), supporting the robustness of the age-limited model. Verification for patients younger than 6 years was not feasible due to limited sample size. These findings suggest that histopathological evaluation using specific tissue markers is likely accurate up to approximately age 7, although further refinement within this critical interval is still required. The diagnosis of DMD does not become more difficult; rather, the evaluation of the disease progression in DMD patients does. This observation implies significant individual variability in the deterioration of muscle condition after the critical interval.

The approach used in this analysis, analyzing the difference between the actual and predicted ABx directly as a variable, is somewhat distinct from standard error analysis methods commonly applied in statistical contexts; typically, statistical error analysis aims to evaluate and improve prediction accuracy or model fit by assessing

indices such as squared residuals, distribution of residuals, and metrics of model accuracy including residual standard error and R-squared statistic (James et al., 2021). In contrast, the present study uniquely employs prediction errors themselves as data for further analysis, rather than simply assessing prediction accuracy. Specifically, the prediction errors (residuals) were used to investigate whether increased errors within certain age groups reflect underlying pathological changes. This approach involved performing variance analysis (Kruskal–Wallis test) on prediction errors across different age intervals. While unconventional from the perspective of conventional error analysis for predictive models, this approach appears to remain scientifically valid. Analyzing residuals as data sources provided insight into appropriate age intervals for subsequent variance analysis, allowing the identification of distinct histopathological changes that would have been difficult to detect using traditional predictive modeling alone. Furthermore, without this residual analysis, defining initial age intervals would have been challenging, thereby limiting the ability to capture true variance in the subsequent analyses. This emphasizes the importance of exploratory analytic steps, even when they initially appear indirect or unconventional.

Isolating the 6–7-year age group revealed that *MFD* appeared nearly static within this interval, whereas others exhibited more pronounced changes, highlighting the significance of treating 6–7 years as a distinct period. Notably, these analyses demonstrated significant differences and large effect sizes in *MFD* between the 1–6 and 6–7 years groups, as well as between the 1–6 and 7–11 years groups (**Table 3**). Additionally, a conventional histopathological parameter, *CFA*, exhibited significant differences between the same age groups, with large effect sizes, aligning with previous observations by Peverelli et al. (Peverelli et al., 2015). Although individual-level data are unavailable, approximate estimates derived from their summary data (Peverelli et al., 2015) suggest a large effect size (e.g., Cohen's *d*: 1.96 [CI 1.19, 2.73]). In addition, while large effect sizes were detected in variance analysis comparing age groups (1–6 vs. 6–7 and 7–11 years), the actual data revealed considerable overlaps in the distribution among these age groups, with some patients older than age 6 showing *CFA* values similar to those younger than age 6. Furthermore, although the initial variance analysis results reported in the main text indicated a significant difference in *CFA* between the 1–6 and 7–11 years age groups, a subsequent review applying false-discovery-rate correction showed that the difference was not statistically significant. This suggests that individual variability in *CFA* tends to widen with increasing age, a factor potentially responsible for large effect sizes observed in group-level analyses.

Such variability might explain the discrepancies and varying interpretations among previous studies (Cardone et al., 2023; Desguerre et al., 2009; Peverelli et al., 2015). This study found considerable inter-individual variability in *CFA*, and this variability is regarded as the principal reason for the discrepancy with the findings of Peverelli et al., who obtained statistical significance by calculating *CFA* through multi-stage averaging during their statistical analysis (Peverelli et al., 2015). Accordingly, the present findings are in line with those reported by Cardone et al., in that no correlation was observed between *CFA* and age (Cardone et al., 2023). Additionally, evaluating *Opaque* and *IntN* as a proportion relative to *MFD* is proposed to provide a more accurate assessment than counting absolute numbers per FOI, as the pathological significance of these fibers likely varies depending on the overall *MFD*. For instance, identifying 10 abnormal myofibers in an FOI with an *MFD* of 1000 should be interpreted differently from observing the same number in an FOI with an *MFD* of 300. Moreover, inconsistent findings across studies (Cardone et al., 2023; Desguerre et al., 2009; Peverelli et al., 2015) suggest limited specificity and reproducibility of these histological parameters. Thus, previous reports including the present study suggest that these conventional histopathological parameters may have inherent limitations, suggesting they might not serve as reliable markers for accurately tracking disease progression of DMD.

Although the specific numerical thresholds identified through logistic regression (approximately 560–563 fibers/mm², **Table 6**) may be dataset-specific due to the intrinsic nature of the analysis in this study, they clearly underscore age 6 as a critical juncture in disease progression. The number of events in the present logistic regression model was 13, satisfying the commonly recommended criterion of at least 10 events per explanatory variable (Peduzzi et al., 1996), thereby mitigating concerns about potential overfitting. However, logistic regression alone could not definitively confirm whether this threshold truly represented the critical age of disease transition or merely reflected the chosen reference age in the logistic model.

In this sense, these analyses were entirely exploratory: the 6–7-year band, although motivated by error analysis profiling, was a post-hoc stratification and should intrinsically be treated as an estimand. Consequently, variance analyses based on that band and the logistic model using age 6 as a fixed reference remain vulnerable to the criticism of arbitrariness. In other words, although correlation analyses served as initial screening for the main analysis, subsequent multiple regression, residual prediction error, variance, age stratification, and logistic regression analyses were strongly exploratory in nature, and from the standpoint of the strict research narrative, they might

be considered close to "negative data." However, key findings emerged through these exploratory analyses, enabling an overall reorganization of the statistical analyses. Consequently, primary and secondary measures were redefined, ultimately providing consistency across the analyses. A particularly important issue in this context was the use of the estimand, critical age, as a reference value.

4.2.2 Confirmatory Statistics

To address these issues, segmented regression analysis was implemented, which provided objective statistical confirmation of the critical breakpoint at approximately 6.25 years, clearly revealing distinct regression slopes before and after this point. As described, the primary and secondary measures were logically reconstructed based on these results. The primary measure was defined as the *MFD* age-breakpoint, while the secondary measures included the statistical power for breakpoint detection, *MFD* cut-offs for age prediction, and the behavior of conventional morphometrics across the age bands delineated by the identified breakpoint. Consequently, the main estimand was explicitly designated as the critical breakpoint age, and the secondary measures naturally followed from this clear definition, providing coherence across all analytical endpoints.

Bayesian methods reinforced the reliability of the results of the segmented regression analysis, demonstrating a clear difference between the rates of *MFD* decline before and after this age threshold. Next, the results of segmented regression analysis needed to be validated. Frequentist methods had inherent limitations, and although bootstrap methods might initially appear suitable, the characteristics inherent to segmented regression analysis necessitate different considerations. Since the primary analysis involves evaluating differences in *MFD* dynamics before and after a breakpoint, randomly resampled datasets generated by bootstrapping must contain a sufficient number of data points both before and after this breakpoint; otherwise, the results of the bootstrapping become unstable. To the author's knowledge, no existing literature has specifically addressed how to validate the results of segmented regression using frequentist approaches. However, analogous guidance exists for logistic regression analysis, which recommends a minimum of 10 events to ensure reliability (Peduzzi et al., 1996). Adopting a similar principle for segmented regression seems reasonable. Yet, implementing a bootstrap procedure with such strict minimum data requirements would likely result in frequent errors and unstable analyses too. Furthermore, restricting resampling in a bootstrap to maintain sufficient data points

contradicts the fundamental assumption of random sampling with replacement, introducing methodological inconsistencies. Overcoming these methodological barriers thus emerged as a significant challenge in segmented regression analysis.

Given these considerations, Bayesian statistical approaches employing probability distributions were deemed more suitable for validation. Monte Carlo-based approaches, of which bootstrapping itself is a form, are more effectively integrated within a Bayesian framework. In this study, two Bayesian methods were employed: one using conjugate posterior distributions with a fixed breakpoint at 6.25 years, and another incorporating weakly informative priors based on the results from segmented regression analysis. While the former approach provides convenient results without requiring Markov Chain Monte Carlo simulation, it cannot estimate the breakpoint itself. This method primarily validates the estimated slopes before and after the breakpoint and assesses the improvement in the segmented regression model over a simple linear regression. Additionally, a second logistic regression analysis using an 80% cutoff for *MFD* values and Monte Carlo simulations estimating misclassification rates across age intervals were conducted. Although these results are summarized earlier, whether these analyses are sufficiently rigorous depends on how extensively the sensitivity analyses of segmented regression should be pursued. Given the inherent limitation that external validation was not feasible, the conclusion was that a more rigorous Bayesian approach capable of estimating breakpoints was also necessary. This approach carries higher computational costs and greater difficulty in setting appropriately weakly informative priors.

Regarding weakly informative priors, the SDs for slopes before and after the breakpoint were set based on doubled standard errors obtained from segmented regression analysis. Additionally, the SD for the breakpoint prior was set at 2 (approximately the width of the 95% CI of the estimated breakpoint), sufficiently capturing uncertainty. Finally, the SD for the intercept prior was taken directly as the observed SD of the *MFD* values. Bayesian analysis results clearly supported the segmented regression findings, including breakpoint estimation. Furthermore, to validate the adequacy of the sample size and enhance the robustness of the present findings, a frequentist Monte Carlo simulation approach was also employed. This approach was chosen because it effectively addresses the inherent limitations of the bootstrap method, specifically, the frequent occurrence of severely unbalanced samples around the breakpoint, as described, by repeatedly generating synthetic datasets and assessing the proportion of successful estimations. In addition, given that the core

analytical approach was segmented regression rather than variance analysis (including effect size calculations), verifying sample size adequacy specifically for segmented regression was essential.

The use of *MFD* values obtained from ATPase-stained slides as the basis for the ROPE analysis was highly effective. Given that the ATPase-stained slides, despite severe deterioration, demonstrated remarkable image recovery through LUT-based digital restoration (**Figure 1A**), their integration into statistical analyses significantly extended their practical utility beyond mere illustrative demonstration. Typically, ROPE thresholds are established by standardizing small-effect heuristics (e.g., ± 0.1 on a standardized scale) (Makowski et al., 2019); however, because external error models were unavailable, a scale-preserving, measurement-based ROPE was preferred here to retain direct interpretability in *MFD* units. Consequently, the ATPase-derived *MFD* values provided an empirical foundation for defining the ROPE, simultaneously addressing both methodological challenges, namely, the lack of previously reported measurement error data and the need to practically utilize severely faded ATPase slides. Thus, this approach is considered a significantly useful application, effectively overcoming critical limitations inherent to the study design and providing robust, evidence-based thresholds for Bayesian validation analyses.

4.2.3 Suggested Histopathological Inflection Point

The present study proposes that age 6 represents a histopathological inflection point in the DMD progression, aligning with a peak in early muscle deterioration as indicated by a marked reduction in *MFD*. Beyond this age, disease progression appears to transition into a more heterogeneous phase, potentially reflecting inter-individual variability in pathological remodeling, which may explain the marked dispersion observed in other conventional parameters such as *MFA* and *CFA*. While the present study posited that the pathophysiology of DMD is driven by recurrent necrosis–regeneration cycles, ultimately leading to an incomplete compensatory process (Nonaka and Nishino, 2021), the findings observed in this study indicate that a substantial decline in *MFD* is already overt between the ages of 1 and 6 years, followed by stabilization at lower levels. This observation suggests that inadequate muscle regeneration begins in the early stages of DMD, contributing to progressive muscle tissue remodeling.

The author firmly believes that the abrupt morphological changes observed in the muscles of DMD patients around the ages of 6–7 years are undeniable. As

described, this study has demonstrated that the *MFD* in DMD patients continues to decline rapidly before the age of 6 years. Possibly, a potential physiological threshold for *MFD* exists. Until this threshold is reached, significant morphological changes in the muscles are not evident. However, once the *MFD* falls below the threshold, muscle homeostasis would be disrupted, leading to rapid and "easily noticeable" morphological changes. This study may imply the existence of such a physiological threshold in *MFD*, which is critical for understanding the progression of DMD. Of course, some patients exhibited *MFD* values above the threshold even after the age of 6, or below the threshold even before the age of 6, indicating that the *MFD* threshold represents a range influenced not only individual differences but also by various research conditions, datasets, sample sizes, and sample quality.

4.3 Clinical Insights and Context of Use

4.3.1 Clinical Relevance

Clinical features of DMD, such as difficulty climbing stairs, a characteristic waddling gait, and frequent falls, typically become evident between approximately 3 and 5 years of age (Birnkrant et al., 2018b; Crisafulli et al., 2020; van Dommelen et al., 2024). In addition, contractures affecting the iliotibial bands, hip flexors, and heel cords generally develop by around age 6 (Brooke et al., 1983). Consistent with these observations, serum CK levels in patients with DMD are significantly elevated from infancy, progressively increasing until around 6 years of age, after which they gradually decline, likely reflecting ongoing muscle deterioration (Zatz et al., 1991). Furthermore, McDonald et al. identified a significant functional decline in the 6-minute walk test performance occurring at approximately age 7 in DMD patients, emphasizing a critical transition in clinical function during this period (McDonald et al., 2013). Similarly, Mayhew et al. reported that steroid-naïve 7-year-old boys with DMD showed notably reduced velocities on timed rise-from-floor and 10-meter walk/run assessments compared with those aged between 4 and 6 years, again underscoring this critical age interval (Mayhew et al., 2022). These clinical findings align closely with the histopathological changes identified in the current study, suggesting that the dynamics of *MFD* could represent an underlying histological basis corresponding to these clinically observed functional transitions.

MFD dynamics may also offer a mechanistic explanation for calf muscle pseudohypertrophy. While pseudohypertrophy has traditionally been attributed to fat and connective tissue infiltration (Walters, 2017), *MFD* dynamics indicate that the

subtle yet homogeneous myofiber hypertrophy observed between ages 1 and 6, in conjunction with stable *Cov*, supports prior evidence suggesting that true myofiber hypertrophy contributes to calf enlargement in the early stages of DMD (Nonaka and Nishino, 2021; Walters, 2017). Actually, Bell and Conen (1967) reported a rapid rise in MFS variability at 5–7 years of age, a histological change that they linked to calf enlargement (Bell and Conen, 1967). This pattern is in agreement with the MFS trajectories described in the current report. These observations point to inverse inflection patterns within the 5–7-year interval and indicate that compensatory myofiber hypertrophy masks its loss until the age of approximately 6 years. After this age, the regenerative reserve would be exhausted and myofiber depletion may be histologically overt.

The observations in this study provide important insights into therapeutic strategies. Fang et al. performed a clustering-based analysis demonstrating that DMD patients with higher steroid-sparing anti-inflammatory agent exposure before 6 years showed slower DMD progression compared with those who initiated it after this same age (Fang et al., 2023). McDonald et al. reported that, among patients aged ≥ 7 years, ambulatory function progressively declined over 48 weeks regardless of steroid treatment status (McDonald et al., 2013). Importantly, they suggested that certain therapeutic agents, such as exon-skipping therapies that aimed to restore dystrophin expression, likely require a sufficient proportion of viable myofibers for optimal efficacy, providing suggestive insights that align with the observed *MFD* dynamics (McDonald et al., 2013). Although a lot of clinical guidelines recommend earlier initiation of steroid treatment depending on individual clinical findings including early functional decline, the optimal timing for clinical intervention remains unknown due to the absence of robust medical evidence (Birnkrant et al., 2018b; Duan et al., 2021; Fontaine Carbonnel et al., 2024; Matthews et al., 2016). Furthermore, concerns regarding potential adverse effects, including growth suppression, impaired bone metabolism, and excessive weight gain, are well-documented (Birnkrant et al., 2018b). Recent randomized controlled trials evaluating vamorolone, a dissociative steroidal anti-inflammatory drug, indicated that steroid-naïve DMD patients treated between 4 and 7 years of age experienced improvements in motor function similar to those achieved with prednisone, but with fewer adverse effects on growth and bone health (Dang et al., 2024; Guglieri et al., 2022), thus supporting the feasibility of earlier therapeutic intervention.

Moreover, multiple therapeutic classes, including corticosteroids, have been

investigated or are currently under evaluation for DMD. Because each class of therapeutic agents acts via a distinct pharmacological mechanism, the impacts of these agents on *MFD* are expected to differ substantially. For instance, myostatin inhibitors such as domagrozumab primarily stimulate muscle hypertrophy (Bogdanovich et al., 2002; Wagner et al., 2020), and are therefore anticipated to influence *MFA* rather than increasing *MFD* itself. In contrast, novel anti-inflammatory treatments, including the NF- κ B inhibitor edasalonexent, might slow the early decline in *MFD* by reducing the extent of fiber necrosis (Finkel et al., 2021). Similarly, exon-skipping approaches and microdystrophin gene therapies, through stabilization of the sarcolemma, could potentially moderate the rate of early *MFD* reduction by decreasing the frequency of necrosis–regeneration cycles (Boehler et al., 2023; Mendell et al., 2020b; Takeda et al., 2021). After the critical period, *MFD* stabilizes at lower levels and gradually declines further, limiting the predictive ability of the present cross-sectional study design regarding treatment efficacy in more advanced disease stages. Consequently, these data do not clarify whether the observed reductions in *MFD* are reversible or whether therapeutic interventions could demonstrate delayed beneficial effects. Additionally, structural and qualitative changes in MFS, driven by continuous cycles of necrosis and regeneration, may follow *MFD* alterations with a temporal delay. These interpretations remain tentative, reflecting well-documented pharmacological insights, available clinical trial results, and the intrinsic detection characteristics of *MFD* identified in this study. Therefore, prospective longitudinal studies are essential to clarify these points further.

4.3.2 Implications for Disease Progression Assessment

This clinical scenario underscores the growing need for reliable biomarkers that can objectively inform the timing and appropriateness of early therapeutic initiation. In this context, *MFD* could serve as an objective, pathology-based biomarker, facilitating more individualized and evidence-based clinical decision-making. For clinical application, establishing baseline *MFD* thresholds in the DMD disease course through future natural history studies would be desirable. However, given the increasing emphasis on earlier corticosteroid and steroid-sparing interventions (Dang et al., 2024; Fang et al., 2023; Guglieri et al., 2022), the recruitment of completely steroid-naïve DMD cohorts for natural history studies will become increasingly challenging from an ethical standpoint. Nonetheless, large-scale validation studies, including both steroid-treated and untreated cohorts, are imperative to further substantiate these findings.

Moreover, whether the low and stable *MFD* dynamics observed after age 6 represent an irreversible state or could potentially be reversed by effective therapeutic interventions remains unclear, as the observations in this study are based solely on steroid-naïve patients. On the other hand, muscle biopsy remains an important tool in contemporary clinical trials, serving as both a primary and secondary outcome measure in the evaluation of novel therapeutic strategies (Boehler et al., 2023; Frank et al., 2020; Mendell et al., 2020a, 2013). Therefore, clinical trials already employing muscle biopsy as an endpoint should incorporate longitudinal assessment of *MFD* dynamics as a monitoring tool alongside conventional motor and/or functional evaluations. Counting-based *MFD* assessment is highly feasible, and currently available digital tools (Dubuisson et al., 2022) capable of automating myofiber enumeration can further facilitate routine clinical implementation. Importantly, in diagnostic pathology, counting-based indices routinely achieve high reproducibility (intraclass correlation coefficients >0.90 for the Ki-67 labelling index and mitotic count (Polley et al., 2015; Scolyer et al., 2003)). Because *MFD* is obtained by complete enumeration of transverse fibers, comparable or superior reproducibility is anticipated in practice. Thus, *MFD* will be expected as a candidate surrogate biomarker.

Several surrogate biomarkers have been employed to support accelerated regulatory approvals for the drug development for DMD; however, each has been subject to ongoing debate regarding their clinical utility and predictive validity. For example, eteplirsen, an exon-skipping therapeutic agent, achieved accelerated approval by the U.S. Food and Drug Administration (FDA) based on its demonstrated ability to increase dystrophin-positive fibers as revealed by immunohistochemistry (Mendell et al., 2013). Nonetheless, this surrogate endpoint has been challenged due to uncertainties in accurately predicting clinical benefit and discrepancies arising from methodological differences in dystrophin quantification (Muntoni et al., 2017; Nature Medicine Editorial, 2016). Similarly, microdystrophin expression has emerged as another candidate surrogate biomarker. While the FDA has recognized micro-dystrophin as a "reasonably likely surrogate," its predictive correlation with meaningful clinical outcomes has yet to be definitively established (Boehler et al., 2023). Thus, two key issues consistently emerge in discussions of surrogate biomarkers: whether they robustly predict clinical outcomes, and whether the observed magnitude of biomarker changes translates into clinically meaningful benefits. Similar challenges are seen when *MFD* is considered to be a potential surrogate biomarker. Specifically, achieving regulatory acceptance of *MFD* would necessitate addressing three key criteria:

analytical validity, clinical validity, and context of use (COU) (Bakker et al., 2022; Goldsack et al., 2020; Ou et al., 2021). Regarding analytical validity, *MFD* measurement is straightforward, objective, and highly reproducible, presenting clear advantages by minimizing subjective interpretation. However, demonstrating clinical validity poses significant challenges. Current data provide preliminary evidence that *MFD* could be clinically relevant, yet direct and longitudinal correlations with functional outcomes remain unexplored. Consequently, establishing robust clinical validity through future prospective studies represents a crucial next step. Additionally, the appropriate COU for *MFD* would likely be early-stage DMD, particularly during the critical 5–7-year age window, in which pathological tissue-level alterations occur. In clinical trials incorporating muscle biopsy endpoints for novel therapeutic interventions, *MFD* could directly quantify tissue-level changes that traditional clinical or functional assessments can only indirectly infer. Nevertheless, its therapeutic effects might initially manifest at the microscopic level within ongoing necrosis–regeneration cycles, with a lag before any measurable changes in the myofiber count become apparent. Hence, improvements in *MFD* alone might not immediately guarantee improvements in muscle quality or functional outcomes. In this sense, the COU proposed here for *MFD* would initially be restricted to detecting early tissue-level alterations in DMD pathology, underscoring the need for careful interpretation and additional validation in future clinical trials.

A noninvasive method to evaluate *MFD* would be highly desirable. Although magnetic resonance imaging can detect relatively pronounced pathological changes such as fat replacement or increase in connective tissue (Finanger et al., 2012; Hollingsworth et al., 2013), its sensitivity in detecting subtle histological alterations at the myofiber level remains uncertain. In addition, the results of the statistical analyses also highlight the limitations of current radiologic imaging analyses in DMD research. While they are noninvasive and beneficial for repeated assessments (Finanger et al., 2012; Hollingsworth et al., 2013; Mathur et al., 2010), they have not yielded significant breakthroughs in terms of understanding muscle pathology. In this study, parameters quantified by "area" measurement did not seem to contribute effectively to evaluating disease progression. Furthermore, the lack of specificity in traditional imaging analyses of *MFA*, necrotic tissue, fat, and connective tissue changes underscores the need for more effective approaches such as radiomics and deep learning (Moreno et al., 2021; Pinheiro et al., 2020; Rinaldi et al., 2023; Shiri et al., 2022). Recently, computed–tomography-based radiomics has shown potential in detecting histologically subtle

features in oncology imaging, such as lung cancer invasiveness, lymphovascular invasion, and other prognostically relevant histopathological features (Cheng et al., 2024; Jiang et al., 2020). In this context, *MFD*, as a histopathological "ground truth", provides a critical reference standard for validating and calibrating these noninvasive image-based biomarkers. In other words, *MFD* may not only serve as a surrogate endpoint itself but also as an essential histopathological reference point to inform and calibrate emerging surrogate biomarkers that await future radiological–histological correlation studies.

4.3.3 Theoretical Clinical–Statistical Integration for Practice

The segmented regression model produced an RMSE of 151.44. Monte Carlo experiments used this value to ask how often a single muscle biopsy in DMD patient would be placed on the wrong side of the 6.25-year threshold. With two million replicates per scenario, the mis-classification rate remained high (approximately 39%) when the true ages differed by only ± 0.25 years, but declined rapidly as the gap widened. At a separation of ± 1 year, the mean error fell below 13%, and at ± 1.25 years, it was below 8%, corresponding to approximately a 92%-rate of correct discrimination between children younger than 6.25 years and those who were 6.25 years or older. Notably, the false-positive and false-negative rates were symmetric across the entire range, indicating that the rule is directionally unbiased.

Because clinical decisions are rarely binary, the second logistic model was also interrogated for the conditional probability of being <6.25 years old at any given *MFD* value. The 20% and 80% posterior probabilities mapped to cut-points of 426.36 and 595.51 fibers/mm², respectively. These translate into three practical zones: a high-probability zone (≥ 595.51 fibers/mm²) in which $\geq 80\%$ of cases are likely <6.25 years old and the therapeutic window may be large; a grey zone (426.36–595.51 fibers/mm²) in which probability hovers around chance and individual clinical judgement is required; and a low-probability zone (≤ 426.36 fibers/mm²) in which $\geq 80\%$ of biopsies come from patients who are ≥ 6 years old and the therapeutic window may be small. However, an important difference between the quadriceps muscle examined in the present study and the biceps muscle typically assessed in the current clinical trials is the normal range of MFS (Walters and Baborie, 2020), necessitating cautious interpretation of these results. Taken together, the analysis indicates that the simple age cut-off is highly reliable for patients younger than 5 years old or older than 7.5 years old. However, within roughly the ± 1.00 -year band around the 6.25-year breakpoint,

clinicians should rely less on a strict dichotomy and more on the *MFD* probabilistic zones, integrated with the broader clinical picture, to guide therapy. Importantly, the optimal way to weigh these *MFD* probability bands against individual clinical factors remains uncertain and should be clarified by future clinicopathological research.

4.4 Limitations and Future Directions

4.4.1 Study Limitations

This study had several limitations. First, because the material consisted of archival specimens and included multiple staining and preparation techniques, full cross-method validation (all stains and frozen vs FFPE) was not feasible. Residual methodological heterogeneity and long-term preservation effects cannot be fully excluded.

Second, the limited sample size, particularly within age strata, may have widened CIs, reduced statistical power, and restricted the complexity of multivariable modelling.

Third, the single-center retrospective design may have introduced selection bias and limits generalizability.

Fourth, the historical nature of the cases meant that detailed longitudinal clinical information was unavailable for many patients, precluding comprehensive clinicopathological correlations.

Finally, although *MFD* is a straightforward measure with high reproducibility and practicality, it inherently simplifies complex histopathological processes in the muscle into a single numerical index, potentially limiting comprehensive pathophysiological interpretations. A larger sample size would have enabled more robust multivariate analyses, potentially revealing additional subtle or interactive pathological factors.

4.4.2 Issues of This Study and Perspectives for Future Research

Despite the inherent constraints associated with archival specimens, the methodological approaches utilized in this study provided a novel perspective on early-stage DMD and its clinical significance.

Reflecting on the intrinsic challenges encountered in this study is essential. This research inherently possesses a different kind of difficulty compared with other studies, arising not only from methodological or analytical complexity but more fundamentally from differences in the "evaluation criteria" and "field-specific

expectations." For instance, in cancer pathology and oncology research, molecular biology approaches such as genetic analysis and targeted therapies are well-established and clearly recognized evaluation standards. Therefore, studies integrating morphological and immunohistochemical assessments with molecular pathological analyses using advanced techniques, such as comprehensive genomic profiling benefit from readily accepted objective validity and well-defined interpretative guidelines, no matter how complex the methods involved may be. In contrast, studies such as the present one, which rely predominantly on quantitative morphometric analyses of archival specimens without the support of predefined analytical frameworks, inherently face unique challenges in gaining immediate recognition and validation within the scientific community.

In the author's prior experience, even single-case studies can achieve international scientific adoption if they clearly align with established evaluation frameworks. Two research projects were independently initiated and completed during the doctoral course. In these studies, the LUT-adjustment technique and Focus Stacking were used to create figures. In 2024, the author reported the first case of morular metaplasia within a high-grade sporadic gastric foveolar-type adenoma (Yamakado et al., 2024). Although this metaplasia had not been previously documented in gastric tumors, the existence of well-established immunohistochemical evaluation criteria in other organs allowed for the credibility and scientific adoption of this rare entity, even in a single-case observation. Similarly, in a separate study, the author described a case of pulmonary adenocarcinoma that spontaneously transformed into large cell neuroendocrine carcinoma with *MYC* amplification in 2025 (Yamakado et al., 2025). Despite the limitation that such a phenomenon inherently requires continuous observation for morphological confirmation, detailed immunohistochemical analysis and comprehensive next-generation sequencing demonstrated objective validity owing to well-documented genome analytical evaluation criteria, strengthening the scientific credibility of the findings. Thus, these two surgical-pathology papers stand as precedents that demonstrate technical proficiency in tackling rare cases, yet they derived their evaluative metrics from pre-existing external frameworks. However, the present study was burdened by the necessity of stepping outside established analytical frameworks, resulting in inherent difficulties regarding scientific validation and adoption. Thus, although all three papers were produced during the doctoral course, they markedly differ in their characteristics.

From a translational research standpoint as well, the present study encountered

several challenges. As described, throughout the project, the following steps had to be undertaken: establishment of methods for data generation, development of quantitative assessment methodologies, selection of appropriate statistical analyses, and formulation of interpretation criteria for the obtained results. However, such an approach can easily appear more subjective, requiring rigorous justification against potential criticism, such as "Why should clinically significant interpretations be derived from such simple counting data?" Thus, this approach relies heavily on the researcher's duty for logical interpretation and persuasive argumentation to establish the scientific validity and significance of the findings. Furthermore, in the context of a chronic progressive disorder such as DMD, demonstrating clinical relevance necessitates evidence of subtle changes and thresholds associated with disease progression. Frequently, studies such as this research do not directly result in clear therapeutic pathways or immediate clinical decision-making. Consequently, developing persuasive arguments that connect these subtle pathological findings to meaningful clinical implications or therapeutic strategies becomes particularly challenging. Accordingly, the challenges pointed out thus far have not been fully resolved within the present study, and subsequent investigations, coupled with external validation and clinical exploration, will be required to establish an objective evaluation framework and secure robust generalizability.

As a future perspective in molecular biological research, based on the clinical implications reported by Fang et al. (Fang et al., 2023) and the critical-period hypothesis proposed in the present study, transcriptome analyses using publicly available Gene Expression Omnibus data represent a promising avenue for elucidating the molecular mechanisms underlying the proposed critical period. In fact, such analyses have already commenced. Integrating multiple human microarray cohorts to increase sample sizes is expected to effectively detect genes differentially expressed before and after the identified critical age. Subsequently, performing network analyses and pathway enrichment analyses may reveal transcriptomic-level evidence corresponding to the morphological and clinical indications of the critical period. Additionally, single-cell RNA sequencing data from dystrophin-deficient mouse models are available within the same dataset. Comparing mild and severe mouse models by simulating critical period transition through pseudobulk and ligand-target network (NicheNet) analyses may further enhance the understanding of this critical period. Although these analyses are planned as subsequent studies, preliminary evidence gathered so far is substantial and already indicates promising potential.

In summary, the intrinsic challenges of this research arose primarily from the

necessity of developing entirely new evaluation criteria rather than relying on established analytical methods or validated biomarkers. Overcoming these challenges led to the identification of the critical age threshold around 6 years of age and the characteristic pattern of *MFD* decline, which may represent important benchmarks for assessing progression of DMD, potentially informing clinical practice. Of course, validation through larger, controlled research remains essential. The author believes that the findings of the present study support the broader utilization of archival samples and contribute to the advancement of future translational research, refining our understanding of the natural history of the DMD and enhancing therapeutic decision-making for this severe, devastating disease.

5. Conclusions

Significant Insights from This Study

1. *MFD* serves as a sensitive biomarker for early-stage DMD progression, revealing a previously unrecognized subclinical phase characterized by accelerated myofiber loss.
2. The identified age threshold of approximately 6 years of age and the observed *MFD* dynamics delineate a critical histopathological transition in disease progression, thereby providing deeper insights into the natural history of DMD.
3. The original semi-quantitative image analysis approach, incorporating novel digital restoration techniques such as LUTs adjustment and Focus Stacking, effectively enabled the detection of a novel biomarker, thus broadening the potential for utilizing archival specimens in translational research.

Directions for Future Research on DMD Pathology

1. Longitudinal and multicenter studies that systematically investigate the influence of early histopathological alterations (e.g., *MFD* dynamics) on long-term disease progression, functional outcomes, and therapeutic responsiveness are needed.
2. Expanded patient cohorts, encompassing both steroid-treated and steroid-naïve individuals, are essential to validate the robustness and clinical utility of *MFD* as a candidate surrogate biomarker.
3. Validating the critical period proposed in the present study through alternative approaches, such as omics analyses, is important for elucidating the pathophysiology of DMD and for advancing therapeutic development.

Potential Issues

1. Recruitment barriers for steroid-naïve patients should be considered. As therapeutic strategies increasingly emphasize early initiation of corticosteroids, steroid-sparing agents, and other emerging drug classes, ethical constraints will complicate enrolling untreated cohorts, thereby restricting future natural history studies.
2. The reversibility of *MFD* alterations is uncertain. Elucidating whether the low, stabilized *MFD* observed after the identified threshold (6 years of age) reflects a definitive, irreversible shift in muscle pathology or if these changes may be responsive to therapeutic interventions remains critical.

Acknowledgements

The author expresses deepest gratitude to Professor Shinya Tanaka (Department of Cancer Pathology, Faculty of Medicine, Hokkaido University) for departmental leadership, constructive guidance, and sustained support. Sincere thanks are extended to Dr. Manabu Ito (Hospital Director, Department of Orthopedics, National Hospital Organization Hokkaido Medical Center) and Dr. Masayoshi Nagao (Former Hospital Director, Department of Pediatrics, National Hospital Organization Hokkaido Medical Center) for their institutional support.

The author also gratefully acknowledges Dr. Yuka Ishikawa (Center Director, Center for Neuromuscular Disease, Child Health and Development, National Hospital Organization Hokkaido Medical Center) for neuromuscular clinical perspectives and continued clinical support, Associate Professor Masumi Tsuda (Department of Cancer Pathology, Faculty of Medicine, Hokkaido University) for constructive advice and general guidance, and Lecturer Zen-ichi Tanei (Department of Surgical Pathology, Hokkaido University Hospital) for neuropathological guidance, sustained discussion, and supervision from study conception through manuscript completion.

Further appreciation is due to Dr. Yusuke Ishida (Department of Pathology, National Hospital Organization Hokkaido Medical Center) for expert diagnostic review and sustained mentorship that formed the foundation of the author's diagnostic training, and to Assistant Professor Hirokazu Sugino (Department of Cancer Pathology, Faculty of Medicine, Hokkaido University) for his considerate support and encouragement.

Moreover, the author thanks the co-authors of the peer-reviewed article designated as the foundation for the present doctoral dissertation for their valuable contributions, including Dr. Koki Ise (Department of Cancer Pathology, Graduate School of Medicine, Hokkaido University), Dr. Taichi Kimura (Department of Pathology, National Hospital Organization Hokkaido Medical Center), Assistant Professor Lei Wang (Institute for Chemical Reaction Design and Discovery [WPI-ICReDD], Hokkaido University), and Assistant Professor Yoshitaka Oda (Department of Cancer Pathology, Faculty of Medicine, Hokkaido University). The members of the Department of Cancer Pathology, Faculty of Medicine, Hokkaido University, and the clinical and laboratory medicine staff at the National Hospital Organization Hokkaido Medical Center are also gratefully acknowledged for their day-to-day support. Regarding language editing, the author acknowledges support from Editage (www.editage.com).

Conflicts of Interest

The author declares no conflicts of interest. No funds were received for this study.

References

- Bakker E, Hendrikse NM, Ehmann F, van der Meer DS, Llinares Garcia J, Vetter T, Starokozhko V and Mol PGM (2022) Biomarker Qualification at the European Medicines Agency: A Review of Biomarker Qualification Procedures From 2008 to 2020. *Clin Pharmacol Ther* 112, 69–80.
- Bell CD and Conen PE (1967) Change in fiber size in Duchenne muscular dystrophy. *Neurology* 17, 902–913.
- Bez Batti Angulski A, Hosny N, Cohen H, Martin AA, Hahn D, Bauer J and Metzger JM (2023) Duchenne muscular dystrophy: disease mechanism and therapeutic strategies. *Front Physiol* 14, 1183101.
- Biggar WD, Skalsky A and McDonald CM (2022) Comparing Deflazacort and Prednisone in Duchenne Muscular Dystrophy. *J Neuromuscul Dis* 9, 463–476.
- Birnkrant DJ, Bushby K, Bann CM, Alman BA, Apkon SD, Blackwell A, Case LE, Cripe L, Hadjiyannakis S, Olson AK, et al (2018a) Diagnosis and management of Duchenne muscular dystrophy, part 2: respiratory, cardiac, bone health, and orthopaedic management. *Lancet Neurol* 17, 347–361.
- Birnkrant DJ, Bushby K, Bann CM, Apkon SD, Blackwell A, Brumbaugh D, Case LE, Clemens PR, Hadjiyannakis S, Pandya S, et al (2018b) Diagnosis and management of Duchenne muscular dystrophy, part 1: diagnosis, and neuromuscular, rehabilitation, endocrine, and gastrointestinal and nutritional management. *Lancet Neurol* 17, 251–267.
- Birnkrant DJ, Bushby K, Bann CM, Apkon SD, Blackwell A, Colvin MK, Cripe L, Herron AR, Kennedy A, Kinnett K, et al (2018c) Diagnosis and management of Duchenne muscular dystrophy, part 3: primary care, emergency management, psychosocial care, and transitions of care across the lifespan. *Lancet Neurol* 17, 445–455.
- Boehler JF, Brown KJ, Beatka M, Gonzalez JP, Donisa Dreghici R, Soustek-Kramer M, McGonigle S, Ganot A, Palmer T, Lowie C, et al (2023) Clinical potential of microdystrophin as a surrogate endpoint. *Neuromuscul Disord* 33, 40–49.
- Bogdanovich S, Krag TOB, Barton ER, Morris LD, Whittemore L-A, Ahima RS and Khurana TS (2002) Functional improvement of dystrophic muscle by myostatin blockade. *Nature* 420, 418–421.
- Brooke MH, Fenichel GM, Griggs RC, Mendell JR, Moxley R, Miller JP and Province MA (1983) Clinical investigation in Duchenne dystrophy: 2. Determination of the

- “power” of therapeutic trials based on the natural history. *Muscle Nerve* 6, 91–103.
- Cardone N, Taglietti V, Baratto S, Kefi K, Periou B, Gitiaux C, Barnerias C, Lafuste P, Pharm FL and Pharm JN (2023) Myopathologic trajectory in Duchenne muscular dystrophy (DMD) reveals lack of regeneration due to senescence in satellite cells. *Acta Neuropathol Commun* 11, 167.
- Cheng DO, Khaw CR, McCabe J, Pennycuik A, Nair A, Moore DA, Janes SM and Jacob J (2024) Predicting histopathological features of aggressiveness in lung cancer using CT radiomics: a systematic review. *Clin Radiol* 79, 681–689.
- Crisafulli S, Sultana J, Fontana A, Salvo F, Messina S and Trifirò G (2020) Global epidemiology of Duchenne muscular dystrophy: an updated systematic review and meta-analysis. *Orphanet J Rare Dis* 15, 141.
- Dang UJ, Damsker JM, Guglieri M, Clemens PR, Perlman SJ, Smith EC, Horrocks I, Finkel RS, Mah JK, Deconinck N, et al (2024) Efficacy and Safety of Vamorolone Over 48 Weeks in Boys With Duchenne Muscular Dystrophy: A Randomized Controlled Trial. *Neurology* 102, e208112.
- Desguerre I, Mayer M, Leturcq F, Barbet J-P, Gherardi RK and Christov C (2009) Endomysial fibrosis in Duchenne muscular dystrophy: a marker of poor outcome associated with macrophage alternative activation. *J Neuropathol Exp Neurol* 68, 762–773.
- Ding X, Li Z, Lin G, Li W and Xue L (2022) Toll-7 promotes tumour growth and invasion in *Drosophila*. *Cell Prolif* 55, e13188.
- Duan D (2018) Systemic AAV Micro-dystrophin Gene Therapy for Duchenne Muscular Dystrophy. *Mol Ther* 26, 2337–2356.
- Duan D, Goemans N, Takeda S, Mercuri E and Aartsma-Rus A (2021) Duchenne muscular dystrophy. *Nat Rev Dis Primers* 7.
- Dubuisson N, Versele R, Planchon C, Selvais CM, Noel L, Abou-Samra M and Davis-López de Carrizosa MA (2022) Histological Methods to Assess Skeletal Muscle Degeneration and Regeneration in Duchenne Muscular Dystrophy. *Int J Mol Sci* 23, 16080.
- Egan KP, Brennan TA and Pignolo RJ (2012) Bone histomorphometry using free and commonly available software. *Histopathology* 61, 1168–1173.
- Elangkovan N and Dickson G (2021) Gene Therapy for Duchenne Muscular Dystrophy. *J Neuromuscul Dis* 8, S303–S316.
- Erkut E and Yokota T (2022) CRISPR Therapeutics for Duchenne Muscular Dystrophy. *Int J Mol Sci* 23.

- Fang Y, McDonald CM, Clemens PR, Gordish H-D, Illei K, Hoffman EP, CINRG DNHS and Vamorolone 002/003/LTE Investigators, and Dang UJ (2023) Modeling Early Heterogeneous Rates of Progression in Boys with Duchenne Muscular Dystrophy. *J Neuromuscul Dis* 10, 349–364.
- Finanger EL, Russman B, Forbes SC, Rooney WD, Walter GA and Vandenborne K (2012) Use of skeletal muscle MRI in diagnosis and monitoring disease progression in Duchenne muscular dystrophy. *Phys Med Rehabil Clin N Am* 23, 1–10, ix.
- Finkel RS, McDonald CM, Lee Sweeney H, Finanger E, Neil Knierbein E, Wagner KR, Mathews KD, Marks W, Statland J, Nance J, et al (2021) A Randomized, Double-Blind, Placebo-Controlled, Global Phase 3 Study of Edasalonexent in Pediatric Patients with Duchenne Muscular Dystrophy: Results of the PolarisDMD Trial. *J Neuromuscul Dis* 8, 769–784.
- Fitzgerald S, Wang S, Dai D, Murphree DH, Pandit A, Douglas A, Rizvi A, Kadirvel R, Gilvarry M, McCarthy R, et al (2019) Orbit image analysis machine learning software can be used for the histological quantification of acute ischemic stroke blood clots. *PLoS One* 14, e0225841.
- Fontaine Carbonnel S, Dabaj I, de Montferrand C, Rippert P, Laugel V, De Lucia S, Ravelli C, Seferian A, Ropars J and Cances C (2024) Choice of compound, dosage, and management of side effects for long-term corticosteroid treatment in Duchenne muscular dystrophy: Guidelines from the Neuromuscular Commission of the French Society of Pediatric Neurology. *Arch Pediatr* 31, 410–418.
- Frank DE, Schnell FJ, Akana C, El-Husayni SH, Desjardins CA, Morgan J, Charleston JS, Sardone V, Domingos J, Dickson G, et al (2020) Increased dystrophin production with golodirsen in patients with Duchenne muscular dystrophy. *Neurology* 94, e2270–e2282.
- Gelman A, Carlin J, Stern H, Dunson D, Vehtari A and Rubin D (2025) Bayesian Data Analysis Third edition (updated electronic edition) [WWW Document]. URL <https://sites.stat.columbia.edu/gelman/book/BDA3.pdf> (accessed 5.6.25).
- Goldsack JC, Coravos A, Bakker JP, Bent B, Dowling AV, Fitzer-Attas C, Godfrey A, Godino JG, Gujar N, Izmailova E, et al (2020) Verification, analytical validation, and clinical validation (V3): the foundation of determining fit-for-purpose for Biometric Monitoring Technologies (BioMeTs). *NPJ Digit Med* 3, 55.
- Guglieri M, Clemens PR, Perlman SJ, Smith EC, Horrocks I, Finkel RS, Mah JK, Deconinck N, Goemans N, Haberlova J, et al (2022) Efficacy and Safety of

- Vamorolone vs Placebo and Prednisone Among Boys With Duchenne Muscular Dystrophy: A Randomized Clinical Trial. *JAMA Neurol* 79, 1005–1014.
- Happi Mbakam C, Lamothe G, Tremblay G and Tremblay JP (2022) CRISPR-Cas9 Gene Therapy for Duchenne Muscular Dystrophy. *Neurotherapeutics* 19, 931–941.
- Harrington D, D’Agostino RB, Gatsonis C, Hogan JW, Hunter DJ, Normand S-LT, Drazen JM and Hamel MB (2019) New Guidelines for Statistical Reporting in the Journal. *N Engl J Med* 381, 285–286.
- Hollingsworth KG, Garrood P, Eagle M, Bushby K and Straub V (2013) Magnetic resonance imaging in Duchenne muscular dystrophy: longitudinal assessment of natural history over 18 months. *Muscle Nerve* 48, 586–588.
- James G, Witten D, Hastie T and Tibshirani R (2021) *An Introduction to Statistical Learning*. Springer US, New York, NY.
- Jiang C, Luo Y, Yuan J, You S, Chen Z, Wu M, Wang G and Gong J (2020) CT-based radiomics and machine learning to predict spread through air space in lung adenocarcinoma. *Eur Radiol* 30, 4050–4057.
- Jütte L, Yang Z, Sharma G and Roth B (2022) Focus stacking in non-contact dermoscopy. *Biomed Phys Eng Express* 8.
- Kass RE and Raftery AE (1995) Bayes Factors. *J Am Stat Assoc* 90, 773–795.
- Kass RE and Wasserman L (1995) A Reference Bayesian Test for Nested Hypotheses and its Relationship to the Schwarz Criterion. *J Am Stat Assoc* 90, 928–934.
- Lakens D (2022) Improving Your Statistical Inferences [WWW Document]. Online Publish. URL https://lakens.github.io/statistical_inferences/ (accessed 4.27.25).
- Li DD-U, Arlt J, Tyndall D, Walker R, Richardson J, Stoppa D, Charbon E and Henderson RK (2011) Video-rate fluorescence lifetime imaging camera with CMOS single-photon avalanche diode arrays and high-speed imaging algorithm. *J Biomed Opt* 16, 096012.
- Makowski D, Ben-Shachar MS, Chen SHA and Lüdtke D (2019) Indices of Effect Existence and Significance in the Bayesian Framework. *Front Psychol* 10, 2767.
- Mathur S, Lott DJ, Senesac C, Germain SA, Vohra RS, Sweeney HL, Walter GA and Vandenborne K (2010) Age-related differences in lower-limb muscle cross-sectional area and torque production in boys with Duchenne muscular dystrophy. *Arch Phys Med Rehabil* 91, 1051–1058.
- Matthews E, Brassington R, Kuntzer T, Jichi F and Manzur AY (2016) Corticosteroids for the treatment of Duchenne muscular dystrophy. *Cochrane Database Syst Rev* 2016, CD003725.

- Mayhew AG, Moat D, McDermott MP, Eagle M, Griggs RC, King WM, James MK, Muni-Lofra R, Shillington A, Gregson S, et al (2022) Functional outcome measures in young, steroid-naïve boys with Duchenne muscular dystrophy. *Neuromuscul Disord* 32, 460–467.
- McDonald CM, Henricson EK, Abresch RT, Florence JM, Eagle M, Gappmaier E, Glanzman AM, PTC124-GD-007-DMD Study Group, Spiegel R, Barth J, et al (2013) The 6-minute walk test and other endpoints in Duchenne muscular dystrophy: longitudinal natural history observations over 48 weeks from a multicenter study. *Muscle Nerve* 48, 343–356.
- Mendell JR, Rodino-Klapac LR, Sahenk Z, Roush K, Bird L, Lowes LP, Alfano L, Gomez AM, Lewis S, Kota J, et al (2013) Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol* 74, 637–647.
- Mendell JR, Sahenk Z, Lehman K, Nease C, Lowes LP, Miller NF, Iammarino MA, Alfano LN, Nicholl A, Al-Zaidy S, et al (2020a) Assessment of Systemic Delivery of rAAVrh74.MHCK7.micro-dystrophin in Children With Duchenne Muscular Dystrophy: A Nonrandomized Controlled Trial. *JAMA Neurol* 77, 1122–1131.
- Mendell JR, Sahenk Z, Lehman K, Nease C, Lowes LP, Miller NF, Iammarino MA, Alfano LN, Nicholl A, Al-Zaidy S, et al (2020b) Assessment of Systemic Delivery of rAAVrh74.MHCK7.micro-dystrophin in Children With Duchenne Muscular Dystrophy: A Nonrandomized Controlled Trial. *JAMA Neurol* 77, 1122–1131.
- Min Y-L, Bassel-Duby R and Olson EN (2019) CRISPR Correction of Duchenne Muscular Dystrophy. *Annu Rev Med* 70, 239–255.
- Moreno S, Bonfante M, Zurek E, Cherezov D, Goldgof D, Hall L and Schabath M (2021) A Radiogenomics Ensemble to Predict EGFR and KRAS Mutations in NSCLC. *Tomography* 7, 154–168.
- Muggeo VMR (2003) Estimating regression models with unknown break-points. *Stat Med* 22, 3055–3071.
- Muntoni F, Fletcher S and Wilton S (2017) Response to “Railroading at the FDA”. *Nat Biotechnol* 35, 207–209.
- Nakagawa S and Cuthill IC (2007) Effect size, confidence interval and statistical significance: a practical guide for biologists. *Biol Rev Camb Philos Soc* 82, 591–605.
- Nature Medicine Editorial (2016) Railroading at the FDA. *Nat Med* 22, 1193.
- Nonaka I and Nishino I (2021) *Muscle Pathology for Clinical Practice*, 5th ed. Nihon Iji Shinpo Sha, Tokyo.

- Ou F-S, Michiels S, Shyr Y, Adjei AA and Oberg AL (2021) Biomarker Discovery and Validation: Statistical Considerations. *J Thorac Oncol* 16, 537–545.
- Peduzzi P, Concato J, Kemper E, Holford TR and Feinstein AR (1996) A simulation study of the number of events per variable in logistic regression analysis. *J Clin Epidemiol* 49, 1373–1379.
- Peverelli L, Testolin S, Villa L, D’Amico A, Petrini S, Favero C, Magri F, Morandi L, Mora M, Mongini T, et al (2015) Histologic muscular history in steroid-treated and untreated patients with Duchenne dystrophy. *Neurology* 85, 1886–1893.
- Pinheiro G, Pereira T, Dias C, Freitas C, Hespanhol V, Costa JL, Cunha A and Oliveira HP (2020) Identifying relationships between imaging phenotypes and lung cancer-related mutation status: EGFR and KRAS. *Sci Rep* 10, 3625.
- Polley M-YC, Leung SCY, Gao D, Mastropasqua MG, Zabaglo LA, Bartlett JMS, McShane LM, Enos RA, Badve SS, Bane AL, et al (2015) An international study to increase concordance in Ki67 scoring. *Mod Pathol* 28, 778–786.
- Rinaldi L, Guerini Rocco E, Spitaleri G, Raimondi S, Attili I, Ranghiero A, Cammarata G, Minotti M, Lo Presti G, De Piano F, et al (2023) Association between Contrast-Enhanced Computed Tomography Radiomic Features, Genomic Alterations and Prognosis in Advanced Lung Adenocarcinoma Patients. *Cancers (Basel)* 15, 4553.
- Rouder JN, Speckman PL, Sun D, Morey RD and Iverson G (2009) Bayesian t tests for accepting and rejecting the null hypothesis. *Psychon Bull Rev* 16, 225–237.
- Sanderson MJ, Smith I, Parker I and Bootman MD (2014) Fluorescence microscopy. *Cold Spring Harb Protoc* 2014, pdb.top071795.
- Scolyer RA, Shaw HM, Thompson JF, Li L-XL, Colman MH, Lo SK, McCarthy SW, Palmer AA, Nicoll KD, Dutta B, et al (2003) Interobserver reproducibility of histopathologic prognostic variables in primary cutaneous melanomas. *Am J Surg Pathol* 27, 1571–1576.
- Shiri I, Amini M, Nazari M, Hajianfar G, Haddadi Avval A, Abdollahi H, Oveisi M, Arabi H, Rahmim A and Zaidi H (2022) Impact of feature harmonization on radiogenomics analysis: Prediction of EGFR and KRAS mutations from non-small cell lung cancer PET/CT images. *Comput Biol Med* 142, 105230.
- Shosekai-Kenkyujo (2021) *Kenbikyo Tettei Customize Guide* [Complete Guide to Microscope Customization]. Shosekai-Kenkyujo, Tokyo.
- Sun C, Serra C, Lee G and Wagner KR (2020) Stem cell-based therapies for Duchenne muscular dystrophy. *Exp Neurol* 323, 113086.
- Takeda S, Clemens PR and Hoffman EP (2021) Exon-Skipping in Duchenne Muscular

- Dystrophy. *J Neuromuscul Dis* 8, S343–S358.
- van Dommelen P, van Dijk O, de Wilde JA and Verkerk PH (2024) Short developmental milestone risk assessment tool to identify Duchenne muscular dystrophy in primary care. *Orphanet J Rare Dis* 19, 192.
- Volinsky CT and Raftery AE (2000) Bayesian information criterion for censored survival models. *Biometrics* 56, 256–262.
- Wagner KR, Abdel-Hamid HZ, Mah JK, Campbell C, Guglieri M, Muntoni F, Takeshima Y, McDonald CM, Kostera-Pruszyk A, Karachunski P, et al (2020) Randomized phase 2 trial and open-label extension of domagrozumab in Duchenne muscular dystrophy. *Neuromuscul Disord* 30, 492–502.
- Walters J (2017) Muscle hypertrophy and pseudohypertrophy. *Pract Neurol* 17, 369–379.
- Walters J and Baborie A (2020) Muscle biopsy: what and why and when? *Pract Neurol* 20, 385–395.
- Yamakado T, Sato-Yazawa H, Ishii J, Kashiwagi K, Kimura T, Tanei Z-I, Yazawa T, Ishida Y and Tanaka S (2025) Spontaneous Transformation from Lung Adenocarcinoma to MYC-amplified Large Cell Neuroendocrine Carcinoma. *Pathol Int* 75, 105–113.
- Yamakado T, Tanei Z-I, Ishikawa Y, Kimura T, Ishida Y and Tanaka S (2024) Sporadic gastric foveolar-type adenoma with morular metaplasia. *Pathol Int* 74, 489–492.
- Zatz M, Rapaport D, Vainzof M, Passos-Bueno MR, Bortolini ER, Pavanello R de C and Peres CA (1991) Serum creatine-kinase (CK) and pyruvate-kinase (PK) activities in Duchenne (DMD) as compared with Becker (BMD) muscular dystrophy. *J Neurol Sci* 102, 190–196.
- Zhou Y, Ru GQ, Yan R, Wang MS, Chen MJ, Yu LL and Wang H (2017) An Inexpensive Digital Image Analysis Technique for Liver Fibrosis Quantification in Chronic Hepatitis B Patients. *Ann Hepatol* 16, 881–887.