


ORIGINAL ARTICLE

Automated immunofluorescence analysis for sensitive and precise dystrophin quantification in muscle biopsies

Tatyana A. Vetter¹  | Stefan Nicolau¹ | Adrienne J. Bradley¹ | Emma C. Frair¹ | Kevin M. Flanigan^{1,2}

¹Center for Gene Therapy, Nationwide Children's Hospital, Columbus, OH, USA

²Departments of Pediatrics and Neurology, Ohio State University Wexner Medical Center, Columbus, OH, USA

Correspondence

Kevin M. Flanigan, The Center for Gene Therapy, The Research Institute, Nationwide Children's Hospital, 700 Children's Drive, Columbus, Ohio 43205, USA.
Email: kevin.flanigan@nationwidechildrens.org

Funding information

National Institute of Arthritis and Musculoskeletal and Skin Diseases; National Institute of Neurological Disorders and Stroke; National Institutes of Health (NIH), Grant/Award Numbers: NS085238, P50AR070604

Abstract

Aims: Dystrophin, the protein product of the *DMD* gene, plays a critical role in muscle integrity by stabilising the sarcolemma during contraction and relaxation. The *DMD* gene is vulnerable to a variety of mutations that may cause complete loss, depletion or truncation of the protein, leading to Duchenne and Becker muscular dystrophies. Precise and reproducible dystrophin quantification is essential in characterising *DMD* mutations and evaluating the outcome of efforts to induce dystrophin through gene therapies. Immunofluorescence microscopy offers high sensitivity to low levels of protein expression along with confirmation of localisation, making it a critical component of quantitative dystrophin expression assays.

Methods: We have developed an automated and unbiased approach for precise quantification of dystrophin immunofluorescence in muscle sections. This methodology uses microscope images of whole-tissue sections stained for dystrophin and spectrin to measure dystrophin intensity and the proportion of dystrophin-positive coverage at the sarcolemma of each muscle fibre. To ensure objectivity, the thresholds for dystrophin and spectrin are derived empirically from non-sarcolemmal signal intensity within each tissue section. Furthermore, this approach is readily adaptable for measuring fibre morphology and other tissue markers.

Results: Our method demonstrates the sensitivity and reproducibility of this quantification approach across a wide range of dystrophin expression in both dystrophinopathy patient and healthy control samples, with high inter-operator concordance.

Conclusion: As efforts to restore dystrophin expression in dystrophic muscle bring new potential therapies into clinical trials, this methodology represents a valuable tool for efficient and precise analysis of dystrophin and other muscle markers that reflect treatment efficacy.

KEYWORDS

Becker muscular dystrophy, computer-assisted image analysis, Duchenne muscular dystrophy, dystrophin, fluorescence microscopy, immunofluorescence

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2021 The Authors. *Neuropathology and Applied Neurobiology* published by John Wiley & Sons Ltd on behalf of British Neuropathological Society.

INTRODUCTION

The dystrophinopathies are a group of X-linked muscular dystrophies caused by mutations in the *DMD* gene. *DMD* encodes dystrophin, a key protein for maintenance of sarcolemmal integrity during muscle contraction. Due to the large size of the gene, comprising 2.4 million bases and 79 exons, *DMD* mutations are relatively common and very diverse. Duchenne muscular dystrophy (DMD) represents the most severe end of the dystrophinopathy spectrum and is generally associated with very little residual dystrophin expression, whereas the milder Becker muscular dystrophy (BMD) often involves partial expression of a variety of functional dystrophin isoforms.

Western blotting has generally served as the primary quantitative assay of dystrophin levels, with immunofluorescence microscopy playing an adjunctive role for qualitative confirmation of relative dystrophin levels and localisation. Modern microscopy techniques, however, have benefited from hardware and software advancements that deliver important advantages in pursuit of reliable dystrophin quantification. Most noteworthy among these are exceptional sensitivity to small and highly localised changes in dystrophin signal, the opportunity to assess whole tissue expression on the basis of individual muscle fibres and the ability to measure dystrophin levels in situ without the confounders of extraction or transfer steps.

The development of molecular therapeutics in recent years has brought a renewed interest in dystrophin quantification methods. Currently approved therapeutics restore only 1–6% of normal dystrophin levels by western blot, highlighting the need for accurate and unbiased quantification of low levels of dystrophin expression [1–4]. Although the amount of dystrophin resulting in clinical benefit is unclear—especially among therapeutically expressed dystrophins, which may be only partially functional—levels of 0.5–3.2% of normal as assessed by western blot may drive meaningful differences in phenotype [5, 6]. Controversy prominently surrounded the matter of low-level dystrophin quantification during regulatory approval of eteplirsen, the first exon-skipping antisense oligonucleotide, underscoring the importance of methodological rigour and objectivity in immunofluorescence quantification of trial biomarkers [7–9]. Outside of clinical trials, sensitive and reproducible dystrophin quantification would also be a powerful tool for understanding genotype–phenotype correlations and the clinical benefits associated with various levels of dystrophin expression in dystrophinopathy patients.

Multiple groups have now developed dystrophin immunofluorescence analysis methods that strive for objectivity and sensitivity across different operators and laboratories [10–16]. These methods differ in the use of whole sections or small regions of interest (ROIs), the degree of automation, software platforms and accessibility in terms of cost and learning curve. Some of the approaches have focused on dystrophin intensity metrics, whereas others can also produce positivity measurements. The good reported intra- and inter-laboratory reproducibility of these methods supports the use of immunofluorescence microscopy as a primary basis of dystrophin quantification [17]. Nonetheless, there remain opportunities for improvement with regards to objectivity, utility and accessibility of these approaches.

Key Points

- We developed a novel automated and unbiased immunofluorescence analysis methodology for quantifying dystrophin positivity and intensity in muscle.
- Using this methodology, we quantified a wide range of dystrophin, spectrin, and laminin expression levels across 15 dystrophinopathy muscle samples and 6 healthy control samples.
- The described approach demonstrated good sensitivity and excellent reproducibility between serial tissue sections and between different operators.
- An accessible, robust, and high-throughput dystrophin quantification method is a valuable tool for preclinical and clinical studies in dystrophinopathies.

We have developed a new dystrophin quantification methodology that has a modest learning curve and requires minimal operator input to collect precise measurements of fibre dystrophin expression across entire muscle sections. These measurements include both fibre-level dystrophin intensity at the sarcolemma and the proportion of the fibre sarcolemmal coverage with positive dystrophin signal, hereafter referred to as positivity. Collecting detailed positivity data in this fashion allows for easy calculation of the percentage of dystrophin-positive fibres (PDPF) retrospectively using any chosen positivity cut-off for considering a fibre positive for dystrophin. To facilitate visualisation and auditing of the results, this method also creates detailed and informative heatmap images reflecting dystrophin intensity and positivity for each fibre. In the present study, we have used a large set of dystrophinopathy and healthy control muscle samples to demonstrate that this method displays remarkable sensitivity and reliability between replicates. We further show that the approach can be adapted for quantification of other sarcolemmal markers and muscle fibre morphology. Together, these validation studies support the adoption of the described immunofluorescence quantification method in characterising a wide range of mutations and clinical trial outcomes.

MATERIALS AND METHODS

Muscle samples

We selected 15 archived biopsy tissue samples (labelled A–O) obtained from male dystrophinopathy patients who had undergone biopsies of the vastus lateralis ($n = 13$) or gastrocnemius ($n = 2$) between 2015 and 2020. Most of the biopsies were performed at Nationwide Children's Hospital, and two were received from external institutions. In anticipation of detecting a wide range of dystrophin expression, samples were selected based on a range of clinical severity from patients classified as DMD, BMD or intermediate muscular dystrophy (IMD), using clinical characterisation we have

described elsewhere [18]. The patients harboured a variety of duplication, deletion, nonsense and splice site mutations, resulting in a range of expected dystrophin expression levels (Table 1). Three male (HC1–3) and three female (HC 4–6) healthy control samples were collected from the hamstrings during anterior collateral ligament repair surgery. All subjects (or their parents, for minors) provided written informed consent to the use of their samples, under protocols approved by the Nationwide Children's Hospital Institutional Review Board.

All samples had been frozen in isopentane chilled with liquid nitrogen according to standard techniques [19]. Samples were sectioned on a cryostat at 10- μ m thickness, and at least four serial sections were collected for each round of staining. Slides were stored at -20°C for 1–3 days prior to staining.

Immunofluorescence staining

Immunofluorescence (IF) staining was performed using a cocktail of the following primary antibodies: rabbit polyclonal anti-dystrophin (Abcam ab15277; 1:200), mouse monoclonal anti-spectrin (Leica NCL-Spec1; 1:100) and rat monoclonal anti-laminin (Sigma L0663; 1:400). The following secondary antibodies were used at 1:500 dilution: goat anti-rabbit IgG highly cross-adsorbed antibody with Alexa Fluor Plus 647 (Invitrogen A32733), donkey anti-mouse IgG highly cross-

adsorbed antibody with Alexa Fluor 568 (Invitrogen A10037) and AffiniPure F (ab')₂ fragment donkey anti-rat IgG with Alexa Fluor 488 (Jackson ImmunoResearch 712-546-153). Spectrin and laminin were co-stained as sarcolemmal markers to assess whether the choice of either marker leads to any meaningful differences in quantification. The C-terminal antibody ab15277 was selected to label dystrophin due to its frequent use for staining human and mouse tissue, use in previous quantification studies [12, 14, 15] and high signal-to-noise ratio.

Sections were outlined in PAP pen, incubated in primary antibody solution in phosphate-buffered saline (PBS) for 1 h, washed twice in fresh PBS for 5 min, incubated in secondary antibody solution in PBS for 30 min and washed again twice for 5 min. One section from each slide was incubated in PBS instead of primary antibody solution. Coverslips were mounted using ProLong Gold antifade mountant with DAPI (Invitrogen P36935). All incubations were performed at room temperature, and stained slides were allowed to set for ≥ 30 min before storing at 4°C overnight.

Microscopy

Three stained sections from each sample were imaged at 10x magnification within 36 h of staining (except for sample C, which had unacceptable artefact in one section in Round 1). Whole-section multichannel images were captured using a Nikon Ti2-E motorised

TABLE 1 Sample information

Sample	Sex	Biopsy site	Age at Bx (years)	Phenotype (age at last exam)	Mutation
A	M	Vastus lateralis	7	DMD/IMD (8)	c.4610dupA (p.N1537fs)
B	M	Vastus lateralis	12	IMD/BMD (12)	Duplication of exons 8–13
C	M	Vastus lateralis	6	DMD/IMD (7)	Duplication of exons 3–43
D	M	Vastus lateralis	6	DMD (9)	Duplication of exon 2
E	M	Vastus lateralis	4	DMD (4)	c.9224 + 5G > A
F	M	Vastus lateralis	3	DMD (6)	Deletion of exon 5
G	M	Gastrocnemius	22	BMD (19)	Deletion of exon 44
H	M	Gastrocnemius	18	BMD (19)	Deletion of exon 51
I	M	Vastus lateralis	2	DMD (5)	Deletion of exons 3–29
J	M	Vastus lateralis	7	DMD (10)	c.4600C > T (p.Gln1534X)
K	M	Vastus lateralis	5	DMD (8)	c.186 + 1G > A
L	M	Vastus lateralis	14	BMD (17)	Deletion of exons 3–7
M	M	Vastus lateralis	8	BMD (8)	c.4294C > T (p.Glu1432X)
N	M	Vastus lateralis	8	BMD (10)	Deletion of exons 13–18
O	M	Vastus lateralis	4	BMD (6)	Deletion of exons 13–36
HC1	M	Hamstring	15	HC	--
HC2	M	Hamstring	16	HC	--
HC3	M	Hamstring	15	HC	--
HC4	F	Hamstring	13	HC	--
HC5	F	Hamstring	16	HC	--
HC6	F	Hamstring	14	HC	--

Abbreviations: BMD, Becker muscular dystrophy; DMD, Duchenne muscular dystrophy; IMD, intermediate muscular dystrophy; HC, healthy control.

microscope with a Lumencor SOLA LED light engine (at 50% power), a Hamamatsu ORCA Fusion camera (in 16-bit ultra-quiet mode) and Nikon Plan Achromat Lambda objectives. Final image resolution was 0.64 $\mu\text{m}/\text{pixel}$.

All exposures were chosen by the trained operators based on best microscopy practices of avoiding signal saturation while maximising the useful range of intensity values detected by the camera in healthy control samples. Identical imaging settings were used for all sections within the same round of staining and imaging. Operators selected exposures independently in Rounds 1 and 2, which approximates the degree of variation that is likely to arise between images collected in different laboratories using different microscopy equipment. For Round 3, the imaging operator was required to use the same exposure settings as those used in Round 1.

IF expression analysis

Image quantification was performed using a custom automated analysis process developed in NIS-Elements AR software (v5.30) with the General Analysis 3 module (Nikon). Analysis was performed on the whole tissue area except for a 100- μm border region along the tissue edges. The only step of the analysis requiring operator input was the manual exclusion of significant artefacts (e.g., tissue folds, mounting bubbles, obvious tissue damage, etc.) from the analysed tissue region. All images were processed using rolling ball background correction with a 30- μm radius for the dystrophin channel and a 10- μm radius for the spectrin and laminin channels. Individual muscle fibres were detected using sarcolemmal signal (spectrin or laminin), and measurements of marker positivity and intensity were collected along the perimeter of each muscle fibre (Figure 1), as detailed below. The average combined automated analysis time was <3.5 min per section. A more detailed description of the method is provided in the Supporting Information.

Conventional muscle fibre segmentation

The sarcolemmal marker (either spectrin or laminin) was used to identify muscle fibres by conventional segmentation. A rough automatic threshold was first applied to segment sarcolemmal image regions, and this binary layer was then inverted and filtered by size and shape to identify likely sarcoplasmic regions. The spectrin/laminin signal intensity was measured in these sarcoplasmic regions and used to calculate the precise threshold for final sarcolemmal pixel segmentation for each image. The final sarcolemmal binary layer was lightly processed and inverted, and the resulting binary objects were again filtered by size and shape for final muscle fibre identification.

AI-dependent muscle fibre segmentation

Spectrin was used as the sarcolemmal marker for artificial intelligence (AI) module training and AI-dependent fibre segmentation. A set of six

sample images (A, E, H, K, O and HC1) from Round 1 was selected for AI training, covering a range of disease severity and morphology. One rectangular region measuring 1 \times 2 mm was cropped from the tissue in each image, and an additional region representing mostly background was cropped from the HC1 image. Muscle fibres were manually identified in each image to create the ground truth for training. Using the NIS-Elements module Segment.ai, a convolutional neural network was trained on this training set for 500 iterations, reaching a final training loss of 0.00388. Whole-section images were then processed through fibre segmentation by the trained Segment.ai, and the binary objects generated by the AI were filtered by size and shape using the same filter parameters as conventional segmentation. The AI trained on image crops from Round 1 was successfully used for fibre segmentation from both Rounds 1 and 2.

Dystrophin positivity analysis

Image pixels positive for dystrophin and spectrin/laminin were identified using thresholds derived from sarcoplasmic signal intensities in each image. Fibre dystrophin positivity was calculated as a ratio of the combined length of dystrophin-positive to spectrin/laminin-positive segments within a 4- μm -wide perimeter region around each fibre. The perimeter rings of adjacent fibres were non-overlapping to ensure little to no overflow of dystrophin signal from one fibre to another (Figure S1). The analysis automatically exported the exact % dystrophin-positive perimeter per fibre, a frequency table of the fibre counts in each decile bin (0–10%, 10–20%, etc.), and a heatmap of colour-coded fibre perimeters reflecting the results (Figure 2A). Muscle fibres with $\geq 30\%$ dystrophin-positive perimeter were considered overall positive and used to calculate the PDPF for each tissue section. This criterion is largely consistent with those used in the two most comparable methods that also evaluate fibre positivity [10, 14], while erring on the side of minimising the rate of false positive fibres.

Dystrophin and spectrin/laminin intensity analysis

Raw fibre dystrophin intensity was calculated as the mean signal intensity of the pixels within the 4- μm perimeter region covering the sarcolemma around each fibre. To correct for tissue autofluorescence and low-level nonspecific signal, the mean sarcoplasmic signal intensity in a given tissue section was also measured and subtracted from each fibre's raw sarcolemmal intensity to produce the final fibre intensity value. The median fibre intensity value for each healthy control replicate was then calculated, and the average of the male control median intensities was input as the denominator for quantification of normalised dystrophin intensity for all fibres in all samples. The analysis automatically exported the exact raw and normalised dystrophin intensity per fibre and a heatmap of colour-coded fibre perimeters reflecting the results (Figure 2A). Spectrin and laminin intensities were calculated using the same method.

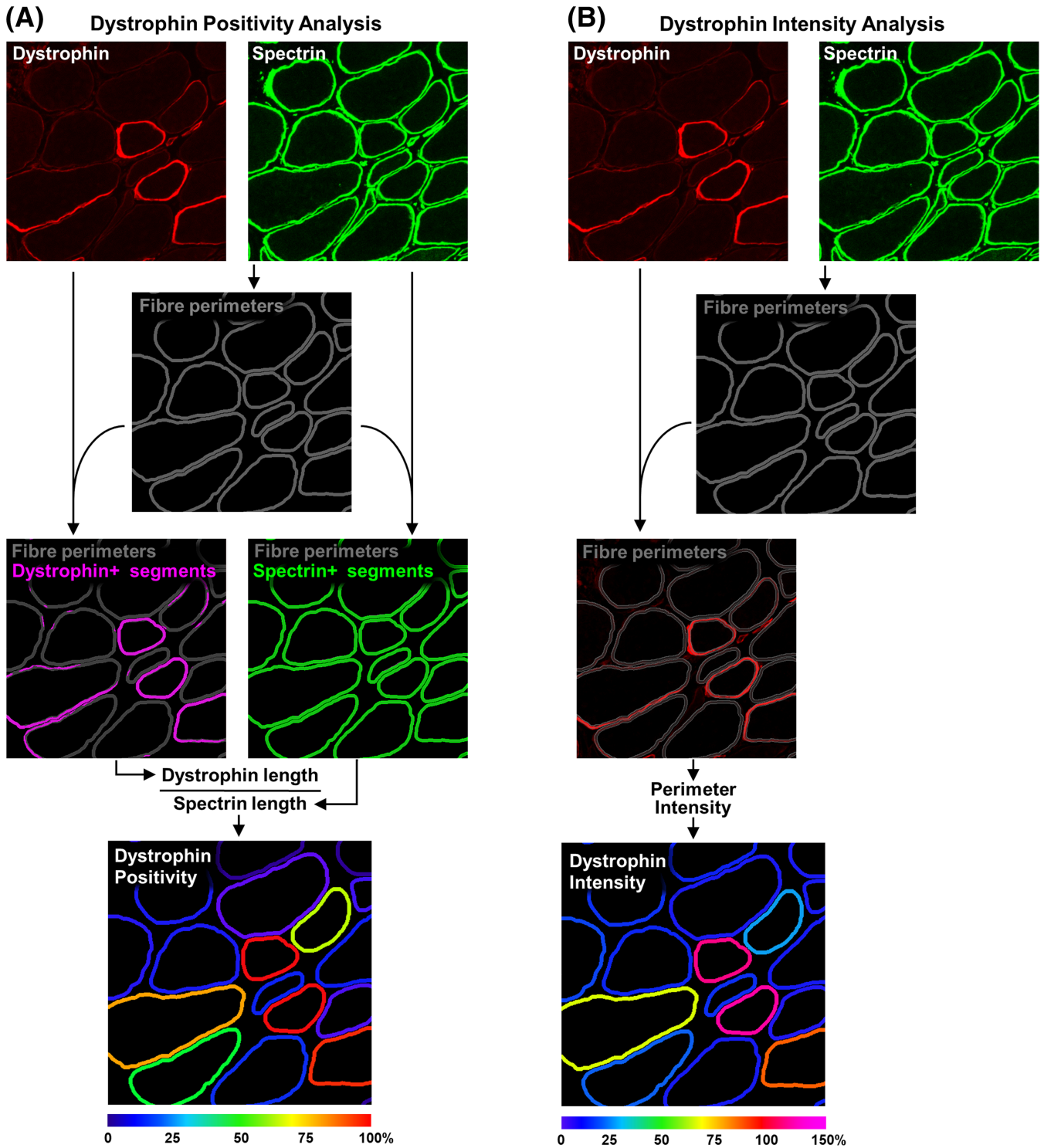


FIGURE 1 Dystrophin immunofluorescence analysis methodology. (A) The dystrophin positivity analysis uses spectrin signal to locate muscle fibre perimeters (grey rings) and identifies the spectrin-positive segments (green) and dystrophin-positive segments (magenta) along the fibre perimeter based on the respective channel images. The combined dystrophin-positive segment length is divided by the spectrin-positive segment length to calculate the dystrophin positivity around the perimeter of each fibre, and the fibre perimeter rings are colour-coded based on the positivity conversion scale to produce the dystrophin positivity heatmap. (B) The dystrophin intensity analysis uses spectrin signal to locate muscle fibre perimeters (grey rings) and measures the mean dystrophin signal intensity for the pixels belonging to each fibre perimeter. The fibre perimeter rings are then colour-coded based on the intensity conversion scale to produce the dystrophin intensity heatmap

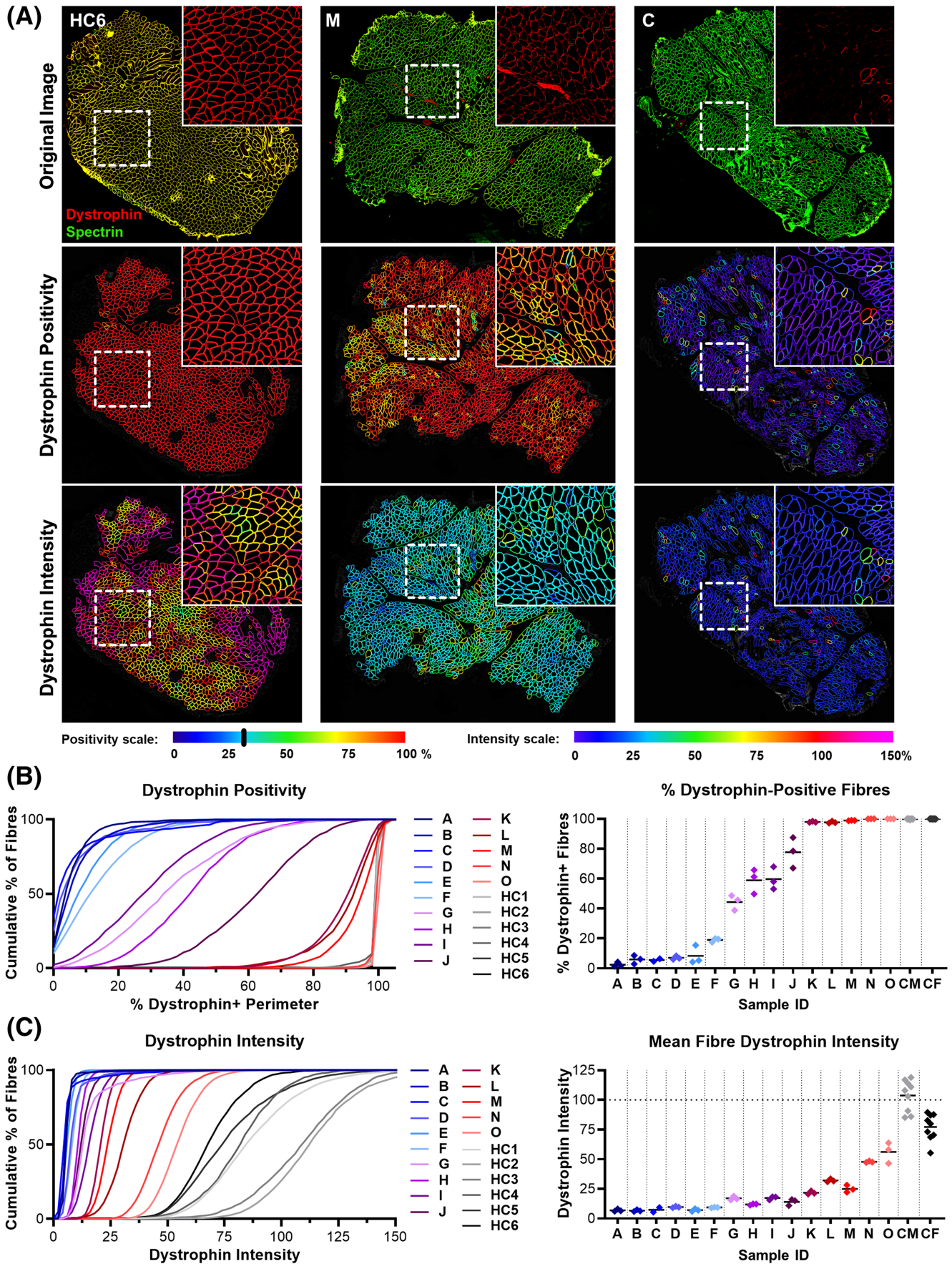


FIGURE 2 Legend on next page.

FIGURE 2 Dystrophin immunofluorescence analysis results. (A) Representative images and matched positivity and intensity heatmaps showing normal (control HC6), intermediate (sample M) and minimal (sample C) dystrophin expression. Original images show merged dystrophin and spectrin, and insets (1 mm²) show only dystrophin. Colour scales indicate the conversions between heatmap colours and fibre dystrophin positivity or intensity. Muscle fibres with $\geq 30\%$ dystrophin-positive perimeter are considered overall positive. (B) Cumulative histogram traces (left) display the fibre positivity profile of each sample. The dotted line at 30% dystrophin-positive perimeter corresponds to the criterion for overall fibre positivity. The total percent dystrophin-positive fibres is plotted for each of three serial sections per sample (right). Solid lines reflect the mean for each sample. All triplicates for three male and three female healthy controls were pooled into control male (CM) and control female (CF) categories. (C) Cumulative histogram traces (left) display the fibre dystrophin intensity profile of each sample. The mean fibre dystrophin intensity, normalised to healthy control intensities, is plotted for each of three serial sections per sample (right). Solid lines reflect the mean of the triplicates for each sample. The horizontal dotted line at 100% reflects the average of the median fibre intensities for male healthy controls. All triplicates for three male and three female healthy controls were pooled into control male (CM) and control female (CF) categories

Inter-operator reproducibility study design

To assess reproducibility, different operators performed IF staining, microscopy and image analysis in three rounds of the study (Figure 3A). For Rounds 1 and 2, operators A and B independently selected imaging exposures. For Round 3, operator A captured images using the exposures selected by operator B in Round 1. Each analysis operator performed quantification on images captured by the other operator, and operator A also repeated the analysis of Round 1 images to assess reproducibility of the method for two operators using an identical image set.

Comparison to prior dystrophin quantification method

The dystrophin image quantification methodology previously developed by our laboratory for ROIs captured via confocal microscopy [15] was adapted for use with widefield whole-section images based on the original description of the method. The method (hereafter referred to as the ‘adapted Taylor method’) was implemented in NIS-Elements AR software and applied to Round 1 images using spectrin as the sarcolemmal marker and an automatic threshold for creating the sarcolemmal mask to ensure objectivity.

Muscle fibre size analysis

Fibre size quantification was automatically performed alongside dystrophin expression quantification during the Dystrophin Positivity Analysis. Minimum Feret’s diameter was used as the primary fibre size metric to minimise the impact of oblique fibre orientation. The fibre size results were automatically exported during the analysis for each fibre, and mean fibre size and the coefficient of variation (CV) were then calculated for each sample.

Data and statistics

All analysis results were automatically exported as tables of numeric data and heatmap images with binary layers overlaid on original image channels. Custom Python scripts were used to automate aggregation

and organisation of numeric data and to facilitate calculation of the percent dystrophin-positive fibres, but such scripts are not required to complete data processing. All data presented are collected from Round 1 images analysed using conventional fibre segmentation based on spectrin, unless otherwise noted.

GraphPad Prism software was used for graphs, descriptive statistics and statistical comparisons. All pairwise comparisons were made using the non-parametric Mann–Whitney test due a small sample size of healthy controls.

RESULTS

Dystrophin quantification

As expected, patient samples showed a wide range of dystrophin positivity and intensity levels (Figure 2). Using a criterion of $\geq 30\%$ positive perimeter for classifying a fibre as positive, all controls had $\sim 100\%$ PDPF. Among dystrophinopathy samples, the PDPF ranged from 1% to 100%, with five samples having a PDPF $< 10\%$ and five others having a PDPF $> 95\%$ (Figure 2B). In contrast, the mean fibre dystrophin intensity in the dystrophinopathy samples was diminished. Because mean fibre dystrophin intensity was 35% higher in male than female controls, patient data were normalised to male controls only. Normalised mean fibre dystrophin intensity ranged from 6.9% to 56% of controls (Figure 2C). Suggesting the sensitivity of this method to detect low-level expressing fibres, all samples with intensity $> 20\%$ had nearly 100% PDPF. Only a single dystrophinopathy sample showed an intensity above 50%, demonstrating significant reduction of dystrophin expression even in samples with nearly 100% PDPF (Figure 2A, middle column).

Inter-operator reproducibility

Inter-operator reproducibility was good for both PDPF and intensity analysis ($R^2 \geq 0.94$ for all comparisons, Figure 3B,C). The exposures chosen by operators A and B in Rounds 1 and 2 were 2.7-fold different for dystrophin and 1.3-fold different for spectrin. The choice of imaging exposure did not affect the reproducibility of the intensity analysis, but the shorter exposure in Round 2 did result in a lower

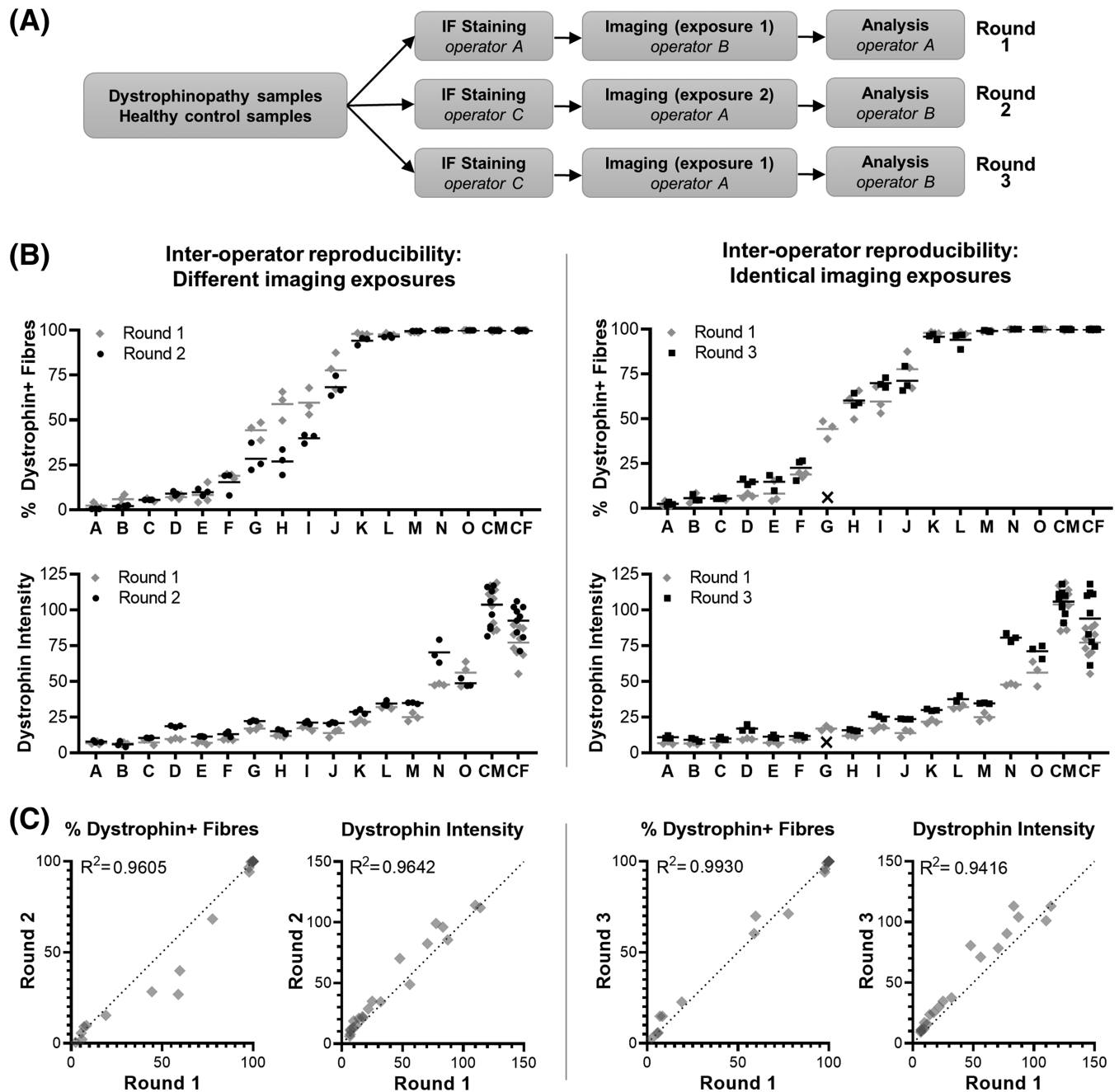


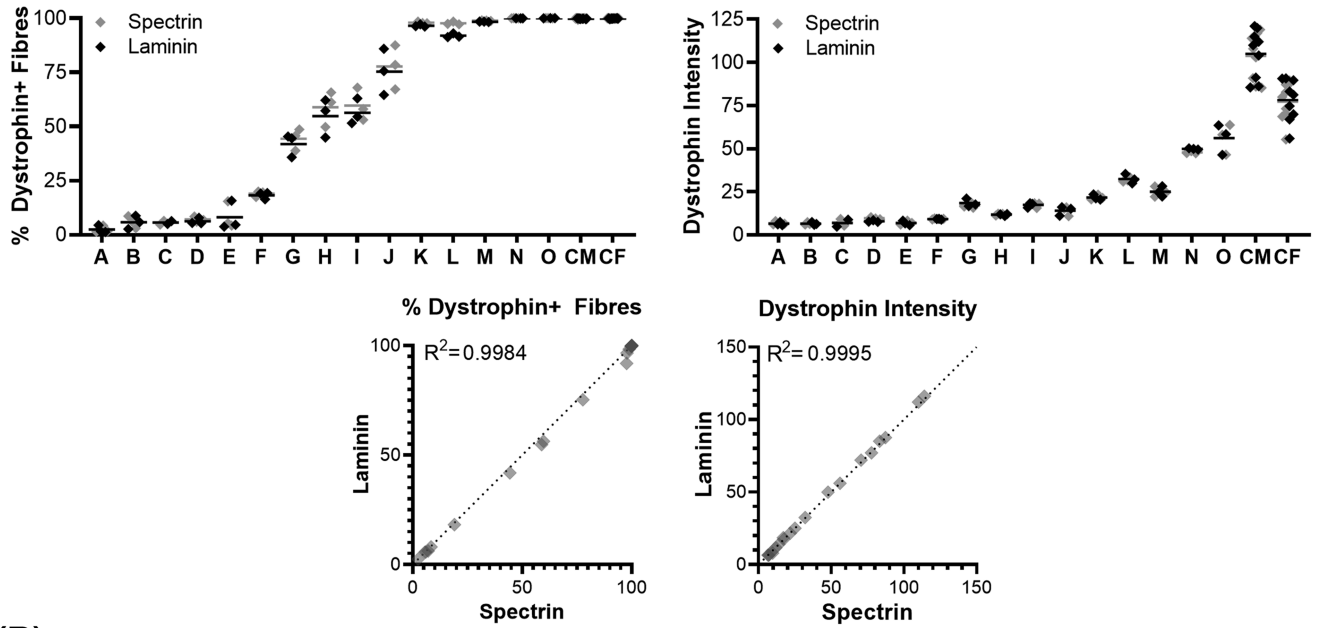
FIGURE 3 Inter-operator reproducibility of the dystrophin analysis. (A) All samples were stained, captured and analysed in three separate rounds to compare results from different operators. In Rounds 1 and 2, images were captured by different operators using different exposure times (selected independently by each operator). In Round 3, images were captured by a different operator than Round 1 but using the same exposure times as Round 1. (B) Percent dystrophin-positive fibres and the mean fibre dystrophin intensity from Rounds 1 vs 2 (left) and Rounds 1 vs 3 (right) are shown for all samples. Round 1 data (grey) are reproduced from Figure 2. Individual points represent each replicate, and the solid horizontal lines represent the replicate mean for each sample and round. All replicates for three male and three female healthy controls were pooled into control male (CM) and control female (CF) categories. The black X symbol indicates the sample that was too depleted for use in Round 3. (C) Correlation plots reflect the concordance between sample dystrophin-positive fibres and fibre dystrophin intensity from Rounds 1 vs 2 (left) and Rounds 1 vs 3 (right), displayed as the triplicate mean for each sample. Healthy control samples are each represented separately on the correlation plots. The diagonal dotted line is the line of identity

PDPF in four samples with intermediate PDPF values (Figure 3B). A secondary analysis of Round 1 images by an additional operator confirmed near-perfect agreement between results from different operators quantifying the same image set with $R^2 = 0.9999$ (Figure S2).

Fibre segmentation methods

Spectrin- and laminin-based fibre segmentation produced highly consistent results for both PDPF and intensity ($R^2 \geq 0.998$ for both

(A)



(B)

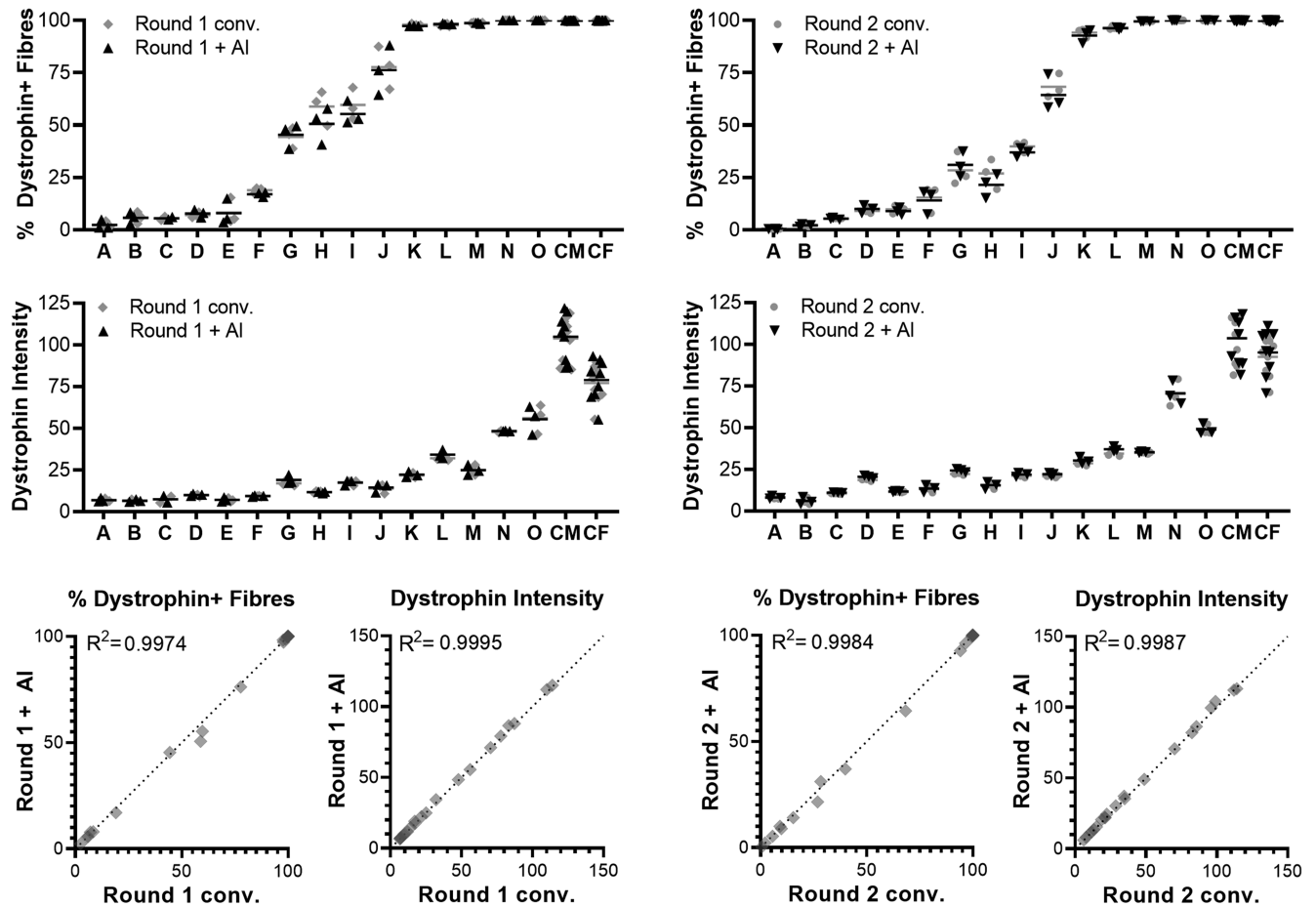


FIGURE 4 Legend on next page.

FIGURE 4 Dystrophin analysis reproducibility between sarcolemmal markers and fibre segmentation methods. (A) All samples were analysed using either the spectrin or laminin channel as the sarcolemmal marker for muscle fibre segmentation and positivity normalisation, performing all other steps identically. The percent dystrophin-positive fibres and the mean fibre dystrophin intensity results from spectrin- vs laminin-based quantification are shown for all samples. The spectrin dataset (grey) is reproduced from Figure 2. (B) All samples were analysed using the spectrin channel to identify muscle fibres by either conventional segmentation methods or by artificial intelligence (AI)-based segmentation, performing all other steps identically. The AI was trained using a small set of regions from Round 1 images, then used to quantify both Rounds 1 (left) and 2 (right). The percent dystrophin-positive fibres and the mean fibre dystrophin intensity quantified using conventional spectrin-based segmentation (Round 1/2 conv.) vs AI-dependent segmentation (Round 1/2 + AI) are shown for all samples. Round 1 conv. and Round 2 conv. datasets (grey) are reproduced from Figures 2 and 3. In grouped plots, individual points represent each replicate, solid horizontal lines represent the mean for each sample and all replicates for three male and three female healthy controls were pooled into control male (CM) and control female (CF) categories. Correlation plots show the triplicate mean for each sample, and all healthy controls are represented individually

comparisons, Figure 4A). However, both metrics were slightly increased with spectrin-based segmentation in most samples (PDPF by 3.5% and intensity by 2%). Conventional and AI-based segmentation methods yielded nearly identical results and AI-based segmentation performed equally well on Rounds 1 and 2 ($R^2 \geq 0.997$ for all comparisons, Figure 4B). Comparisons between spectrin-based conventional segmentation and laminin-based conventional segmentation or AI-dependent segmentation did not reveal any meaningful differences in total fibre counts identified by any of the methods (Figure S3).

Consistency with prior dystrophin quantification methods

A comparison of the current analysis to an adaptation of a previous dystrophin quantification method described by our group [15] was performed to ensure that the new methodology performs consistently with earlier approaches. Because the prior methodology is only able to measure overall dystrophin fluorescence intensity at the sarcolemma, we compared the mean fibre dystrophin intensities obtained using the new methodology to mean sarcolemmal intensities measured using the adapted Taylor method. The comparison demonstrates excellent concordance between the results, with an R^2 value of 0.979, confirming strong consistency between the two methods when applied to identical images (Figure S4).

Sarcolemmal markers

The mean intensity of sarcolemmal markers varied significantly across dystrophinopathy samples, ranging from 60% to 250% of healthy controls (Figure 5A,D). The intensity of the two markers was correlated (Figure 5B), and both were significantly increased in dystrophinopathy samples compared with the combined male and female control group (167% of control for spectrin and 140% for laminin, $p < 0.05$, Figure 5C). The small number of male controls precluded a comparison to male controls only.

Among healthy controls, there was also variation in the intensity of dystrophin, spectrin and laminin (Figure 6). This variation was seen between individual fibres within each sample, with an interquartile

range of 26–52%, as well as between samples, with individual sample means ranging from 63% to 125% of the average of male control median intensities. All three sarcolemmal markers varied to similar degrees and correlated within individual samples. Sarcolemmal marker intensity was higher in males than females, and age is unlikely to explain these differences due to high similarity within and between male and female control groups.

Fibre morphology

Mean muscle fibre sizes were smaller in dystrophinopathy samples than in healthy controls (Figure 7A). While there was an overall correlation between age and fibre size, the four dystrophinopathy samples aged ≥ 12 also showed a mean fibre diameter of 79% of controls ($p < 0.05$). The mean CV of fibre size was increased to 42% in dystrophinopathy samples, compared with 28% in controls ($p < 0.01$, Figure 7B,D). The degree of variation in fibre size did not correlate with dystrophin expression. Conventional and AI-based segmentation methods produced very similar results, but in all samples, fibre size variability was slightly reduced with AI-based segmentation (Figure 7C).

DISCUSSION

Unbiased and reliable dystrophin quantification is a high priority for preclinical studies and clinical trials aiming to restore dystrophin expression. Past controversy related to dystrophin immunofluorescence quantification [8, 9] and relative insensitivity of western blotting at low expression levels [15, 17] underscore the importance of accessible and objective methods for evaluating baseline dystrophin expression and therapeutic efficacy. Although use of the more sensitive capillary immunoassay as an alternative to conventional western blotting may provide the desired improvements in sensitivity [20, 21], microscopy remains critical for assessing cellular localisation and tissue distribution of markers. Our results demonstrate that the muscle IF analysis developed on the NIS-Elements software platform can provide sensitive and reproducible dystrophin quantification across a large range of expression levels. Importantly, this method is highly objective and largely automated, allowing for high-throughput analysis

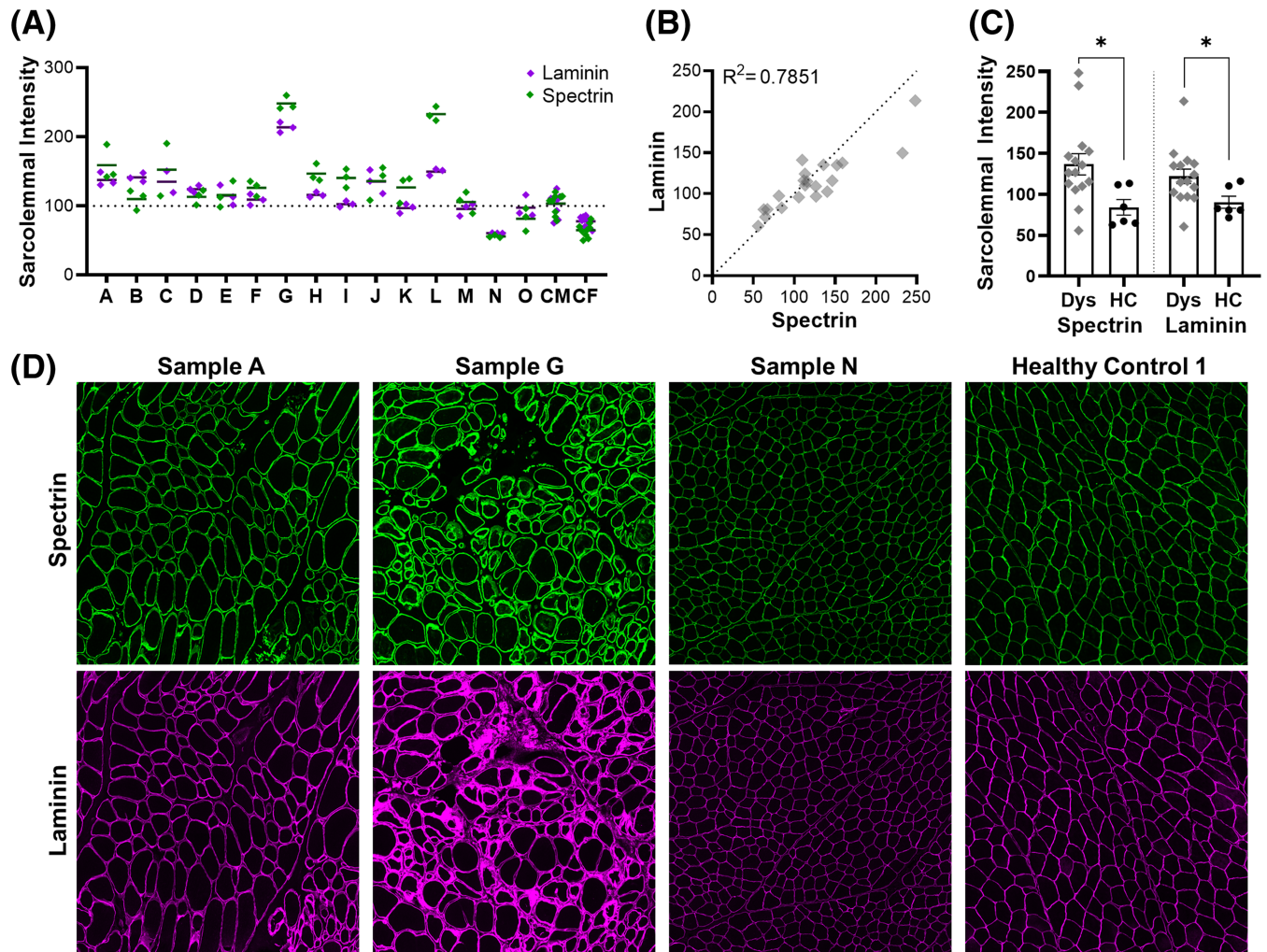


FIGURE 5 Quantification of sarcolemmal marker immunofluorescence intensities. (A) The mean fibre spectrin (green) and laminin (purple) intensities, normalised to healthy control intensities, are plotted for each of three serial sections per sample. Solid lines reflect the mean of the triplicates for each sample. All triplicates for three male and three female healthy controls were pooled into control male (CM) and control female (CF) categories. The horizontal dotted line at 100% reflects the average of the median fibre intensities for male healthy controls. (B) Concordance between mean fibre spectrin and laminin intensities, displayed as the triplicate mean for each sample. Healthy control samples are each represented separately. The diagonal dotted line is the line of identity. (C) Comparison of mean fibre spectrin and laminin intensities between dystrophinopathy samples (Dys) and healthy controls (HC), based on the triplicate mean for each sample. Points represent individual samples, and the bars reflect the mean \pm standard error. Groups compared using Mann–Whitney test; $*p < 0.05$. (D) Representative images of spectrin (green) and laminin (magenta) signal in a typical Duchenne muscular dystrophy (DMD) sample (A), the highest-expressing sample (G), the lowest-expressing sample (N) and a healthy control. Image region size = 1 mm²

of study samples by personnel not formally trained in muscle pathology.

In this study, serial sections showed high consistency for PDPF and dystrophin intensity values, indicating strong intra-operator repeatability. However, the dystrophinopathy samples with intermediate PDPF showed the greatest divergence between triplicates in all rounds, and PDPF showed greater variation between triplicates than intensity measurements. This likely reflects the binary nature of overall fibre positivity, leading to high sensitivity of this measure to small variations in signal intensities near the threshold for pixel positivity, especially in samples where many fibres are also close to the cut-off for fibre positivity. In addition, the mean sample PDPF diverged only

for intermediate dystrophinopathy samples between Rounds 1 and 2, possibly due to the effects of different signal-to-noise ratios in images from these rounds on analysis parameters.

This observation highlights the need to define robust standards and benchmarks for imaging settings that can be generalised across different instruments and operators when porting this assay between laboratories. Nevertheless, reproducibility of dystrophin intensity results between Rounds 1 and 3, which used identical imaging settings, was not better than between Rounds 1 and 2, with different exposures. This comparison suggests that variation in staining may have been a stronger source of variability in intensity quantification than the choice of imaging instruments and settings. This underscores

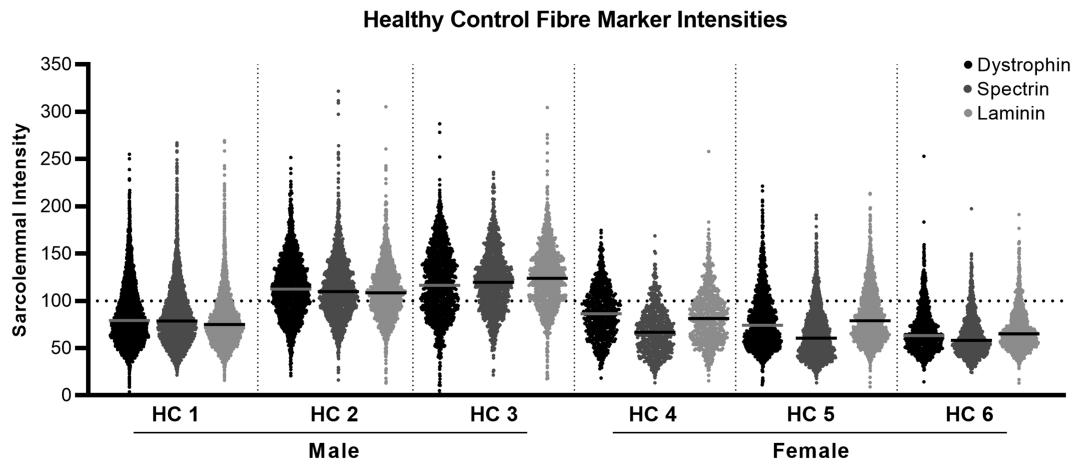


FIGURE 6 Quantification of healthy control signal intensities. Fibre dystrophin (black), spectrin (dark grey) and laminin (light grey) normalised intensities are plotted for all healthy controls. The median replicate is plotted for each control sample, with each point representing an individual muscle fibre in that sample image. The solid lines represent the median fibre intensity for each sample and marker. The horizontal dotted line at $y = 100$ reflects the average of the median fibre intensities in male healthy controls

the importance of optimising and standardising tissue processing and staining protocols to effectively reduce assay variability. Future work will focus on validating the consistency of this quantification approach using images of sections from the same samples captured on different instruments at separate sites.

AI tools are becoming more common and accessible within image processing and analysis software [22], and future investigators will be able to choose between conventional and AI-based approaches. Although conventional segmentation methods are more transparent and tuneable than an AI tool, a properly trained AI may perform better in samples with poorer tissue or staining quality and more interrupted sarcolemmal perimeters. Our results confirmed that the analysis produces very consistent data using either conventional segmentation or a deep-learning-based segmentation AI tool in NIS-Elements. The AI-based approach performed equally well on tissue from Round 1, from which the training dataset was derived, and on Round 2, which was imaged at a different exposure. One minor but consistent difference between results from different segmentation methods was the slightly higher fibre size CV using conventional segmentation. By visual inspection, it appears that the phenomenon of mistakenly identifying spaces between fibres as additional muscle fibres, although very rare using either method, happens slightly more often using conventional segmentation. Despite the infrequent nature of this error, it may have a detectable impact on fibre size CV.

While some previously published methods also report good sensitivity and reproducibility across a range of dystrophin expression levels [17], the approach presented here has additional advantages. Three of the previous methods used ROIs as the basis for sarcolemmal intensity measurements, which were collected hyper-locally [11], on a fibre basis [12] or for all sarcolemmal pixels [15]. Although these methods are highly accessible, use of ROIs for quantification risks random or systematic bias in the selection of the ROIs that is avoided using whole-tissue analyses. Additionally, inability to measure fibre positivity with these analyses is at odds with the frequent preference

for using PDPF as a primary dystrophin expression metric in many studies. However, we have confirmed excellent consistency between mean sarcolemmal dystrophin intensity results obtained using an adaptation of the method presented by Taylor et al. [15] and mean fibre dystrophin intensities measured using our new methodology. Recently, two methods have been described that can perform quantification of whole tissue sections and measure both dystrophin positivity and intensity of each fibre [10, 14]. The Flagship MuscleMap approach employs manually selected thresholds and a proprietary tuneable algorithm to achieve this [10], while another method utilising Definiens software automatically derives thresholds for dystrophin-positive pixels from sarcoplasmic intensities, similarly to the analysis described here [14]. Despite the considerable strengths of this Definiens-based approach, it remains less accessible than NIS-Elements by virtue of its cost and learning curve. The present method also provides greater detail in the visual heatmaps of dystrophin positivity and intensity results on a continuous colour scale, in contrast to categorical annotations provided by the Flagship and Definiens approaches. Although published dystrophin quantification methods were validated using a variety of dystrophin antibodies and different modes of image acquisition, these aspects are not expected to have a profound impact on a carefully designed method.

Both laminin and spectrin have been used to define muscle fibres across published studies, and both markers showed very consistent results in our analysis and good performance for fibre segmentation even in potentially problematic tissue regions (Figure S5). However, in the minority of samples where the spectrin- and laminin-based PDPF did diverge more noticeably, the spectrin-based PDPF was always higher. This trend may be related to the fact that dystrophin and spectrin share the same subsarcolemmal localisation whereas laminin localises extracellularly, but a more specific explanation is not evident based on the available data. Overall, dystrophinopathy samples displayed a wide range of spectrin and laminin intensities that often fell above healthy controls and did not correlate with dystrophin

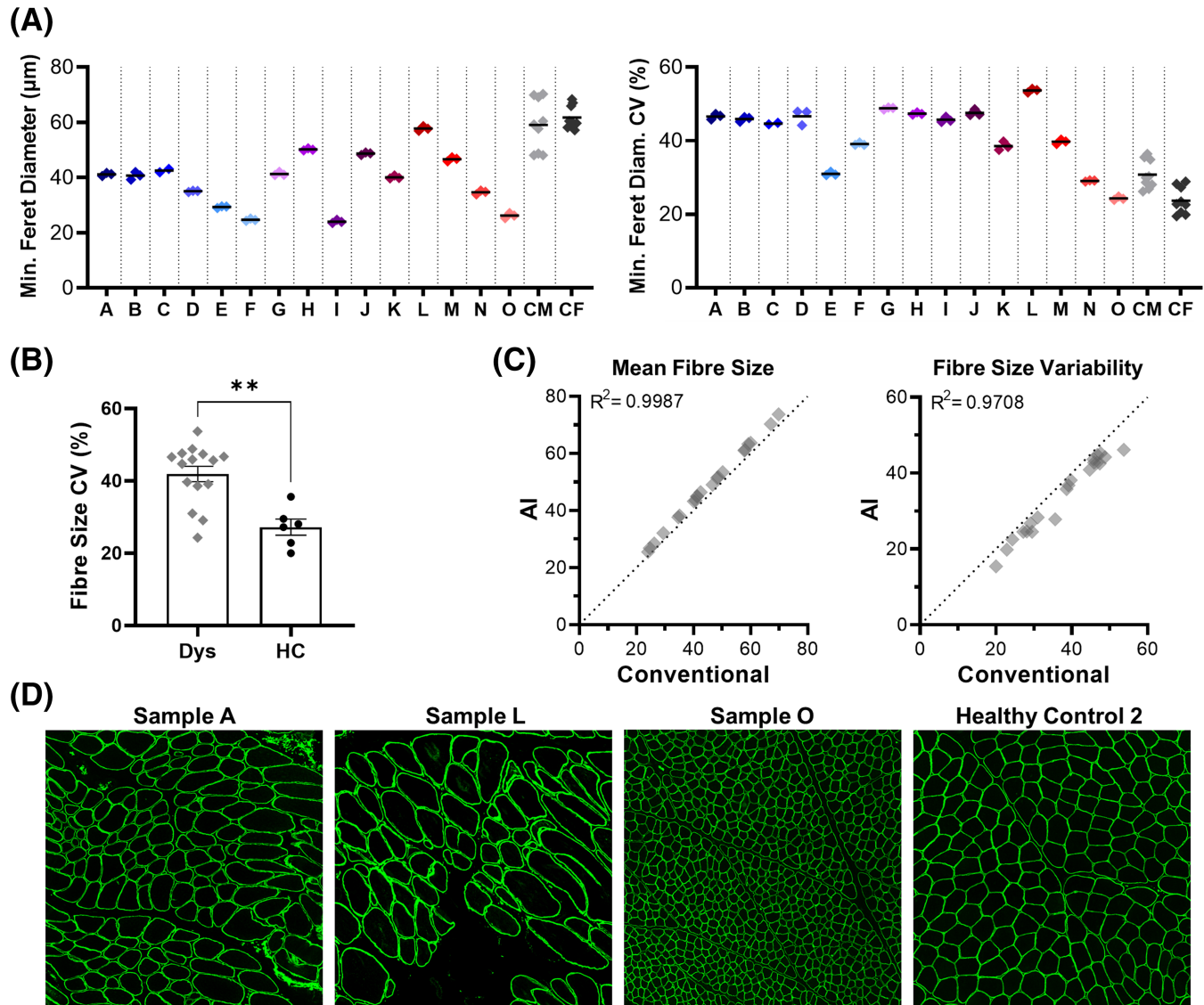


FIGURE 7 Muscle fibre size quantification. (A) Mean muscle fibre size (left) and coefficient of variation (CV; right), measured as the minimum Feret's diameter using the spectrin channel, are plotted for each of three serial sections per sample. Solid lines reflect the mean of the triplicates for each sample. All triplicates for three male and three female healthy controls were pooled into control male (CM) and control female (CF) categories. (B) Comparison of fibre size variability between dystrophinopathy samples (Dys) and healthy controls (HC), based on the triplicate mean for each sample. Points represent individual samples, and the bars reflect the mean \pm standard error. Groups compared using Mann-Whitney test; $**p < 0.01$. (C) Correlation plots reflect the concordance in fibre size mean (left) and CV (right) quantification between conventional fibre segmentation and artificial intelligence (AI)-dependent fibre segmentation for all samples. Correlation plots show the triplicate mean for each sample, and all healthy controls are represented individually. The diagonal dotted line is the line of identity. (D) Representative images of spectrin signal demonstrating fibre size in a typical dystrophinopathy sample (A), the sample with the largest mean fibre size and CV (L), the sample with the smallest mean fibre size and CV (O) and a healthy control. Image region size = 1 mm^2

expression. This finding has been reported previously [13] and contrasts with earlier work by our group that found no significant differences in spectrin between patients and controls [15], which may be explained by differences in the method of quantification and the number of controls.

Additionally, the healthy control samples included in the current study showed a considerable range of fibre dystrophin intensities both within and between samples, as has been shown to varying degrees in previous work [11–13]. These results raise important

considerations regarding normalisation, and suggest that the practices of normalising dystrophin expression to sarcolemmal markers or to only one healthy control may yield misleading results [13], especially with ROI-based methods. While previous studies had reported that normal dystrophin levels do not differ by sex [20], we did find that dystrophin levels in this set of healthy controls appeared higher in males than females. This difference may be age-dependent, and additional experiments with a larger set of healthy control samples of different ages will be needed to characterise dystrophin

expression levels in males and females across the lifespan. We nevertheless recommend careful attention to selection of control tissues, avoiding of the use of histologically normal tissues obtained at clinical biopsy for neuromuscular complaints (as these cannot necessarily be presumed to represent truly normal muscle) and attempting to obtain age- and sex-matched samples whenever possible.

The samples we analysed reproducibly quantified dystrophin expression levels across a wide dynamic range, but we note that levels of expression did not correlate with dystrophinopathy disease classification in all patients (Figure S6). As an example, patient K (classified as DMD) showed 98% PDPF and 23% intensity, whereas patient G (classified as BMD) showed 44% PDPF and 17% intensity. This result is not surprising, as similar findings have been observed in other studies of IF quantification [15, 17]. It is an informative example that highlights the risk of using only PDPF as an outcome measure in therapeutic trials and is likely due to several factors. First, classification is based upon clinical features, including most robustly age at loss of ambulation (LOA); for younger patients, LOA will not have occurred, making clinical classification less reliable. Second, a growing body of evidence supports the importance of genetic modifiers of disease severity, independent of the *DMD* mutation [23–26]. Third, each mutation leads to expression of a different dystrophin protein, which can be expected to have different functional implications. Finally, a reduction in the routine use of muscle biopsy for dystrophinopathy diagnosis has led to an overrepresentation of patients with discordant phenotypes and genotypes among recent biopsies. Importantly, we do not propose that this method can reliably distinguish DMD from BMD—a question that this study was not designed to address—but do propose it is a robust method suitable for measuring fold change of a given dystrophin protein expressed in response to a therapeutic intervention, as we have recently shown in a mouse model [27].

Together, our results provide compelling support for the use of this dystrophin analysis methodology. It is suitable for measuring low levels and small differences in dystrophin expression in the context of different *DMD* mutations, with good repeatability between serial sections and reproducibility between operators. This analytic method also has a relatively mild learning curve and is readily portable between laboratories as a set of algorithm files, requiring relatively little time and effort to launch at a new site. Future validation work will focus on confirming the reproducibility of this method using images captured in different laboratories, and on further investigating dystrophin levels across age and sex in healthy controls. In the coming years, the wider adoption of immunofluorescence quantification methods will help establish the significance of low-level dystrophin expression in clinical practice and will enable more sensitive and detailed comparisons of *DMD*-targeting gene therapies.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Michael Lawlor for helpful discussion. This work was supported in part by NIH grants P50AR070604 from NIAMS and NS085238 from NINDS (to KMF). An abstract with a similar title has been previously published by our group as part of the record of the 2021 congress of the World Muscle Society.

CONFLICT OF INTEREST

The authors have no conflicts of interest to report.

ETHICS STATEMENT

All subjects (or their parents, for minors) provided written informed consent to the use of their samples, under protocols approved by the Nationwide Children's Hospital Institutional Review Board.

AUTHOR CONTRIBUTIONS

KMF and TAV conceived the project, and TAV designed the analysis process. TAV, AJB and ECF designed and performed the experiments. TAV, SN and KMF interpreted the results. TAV and SN processed and analysed the results and drafted the manuscript.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/nan.12785>.

DATA AVAILABILITY STATEMENT

Data published in this manuscript, including software algorithms and scripts, are available from the corresponding author upon reasonable request.

ORCID

Tatyana A. Vetter  <https://orcid.org/0000-0002-9904-9140>

REFERENCES

- Charleston JS, Schnell FJ, Dworzak J, et al. Eteplirsen treatment for Duchenne muscular dystrophy: Exon skipping and dystrophin production. *Neurology*. 2018;90(24):e2146–e2154.
- Komaki H, Takeshima Y, Matsumura T, et al. Viltolarsen in Japanese Duchenne muscular dystrophy patients: A phase 1/2 study. *Ann Clin Transl Neurol*. 2020;7(12):2393–2408.
- Clemens PR, Rao VK, Connolly AM, et al. Safety, tolerability, and efficacy of viltolarsen in boys with Duchenne muscular dystrophy amenable to Exon 53 Skipping: a phase 2 randomized clinical trial. *JAMA Neurol*. 2020;77(8):982–991.
- Frank DE, Schnell FJ, Akana C, et al. Increased dystrophin production with golodirsen in patients with Duchenne muscular dystrophy. *Neurology*. 2020;94(21):e2270–e2282.
- de Feraudy Y, Ben Yaou R, Wahbi K, et al. Very low residual dystrophin quantity is associated with milder dystrophinopathy. *Ann Neurol*. 2021;89(2):280–292.
- Waldrop MA, Gumienny F, el Husayni S, Frank DE, Weiss RB, Flanigan KM. Low-level dystrophin expression attenuating the dystrophinopathy phenotype. *Neuromuscular Disorders: NMD*. 2018; 28(2):116–121.
- Mendell JR, Rodino-Klapac LR, Sahenk Z, et al. Eteplirsen for the treatment of Duchenne muscular dystrophy. *Ann Neurol*. 2013;74(5): 637–647.
- Unger EF, Califf RM. Regarding "Eteplirsen for the treatment of Duchenne muscular dystrophy". *Ann Neurol*. 2017;81(1):162–164.
- Aartsma-Rus A, Arechavala-Gomez V. Why dystrophin quantification is key in the eteplirsen saga. *Nat Rev Neurol*. 2018;14(8): 454–456.
- Aeffner F, Faelan C, Moore SA, et al. Validation of a muscle-specific tissue image analysis tool for quantitative assessment of dystrophin staining in frozen muscle biopsies. *Arch Pathol Lab Med*. 2019;143(2): 197–205.

11. Arechavala-Gomez V, Kinali M, Feng L, et al. Immunohistological intensity measurements as a tool to assess sarcolemma-associated protein expression. *Neuropathol Appl Neurobiol.* 2010;36(4):265-274.
12. Beekman C, Sipkens JA, Testerink J, et al. A sensitive, reproducible and objective immunofluorescence analysis method of dystrophin in individual fibers in samples from patients with duchenne muscular dystrophy. *PLoS ONE.* 2014;9.
13. Sardone V, Ellis M, Torelli S, et al. A novel high-throughput immunofluorescence analysis method for quantifying dystrophin intensity in entire transverse sections of Duchenne muscular dystrophy muscle biopsy samples. *PLoS ONE.* 2018;13(3):1-21.
14. Scaglioni D, Ellis M, Catapano F, et al. A high-throughput digital script for multiplexed immunofluorescent analysis and quantification of sarcolemmal and sarcomeric proteins in muscular dystrophies. *Acta Neuropathol Commun.* 2020;8(1):1-16.
15. Taylor LE, Kaminoh YJ, Rodesch CK, Flanigan KM. Quantification of dystrophin immunofluorescence in dystrophinopathy muscle specimens. *Neuropathol Appl Neurobiol.* 2012;38(6):591-601.
16. Aartsma-Rus A, Morgan J, Lonkar P, et al. Report of a TREAT-NMD/World duchenne organisation meeting on dystrophin quantification methodology. *Journal of Neuromuscular Diseases.* 2019;6(1):147-159.
17. Anthony K, Arechavala-Gomez V, Taylor LE, et al. Dystrophin quantification: Biological and translational research implications. *Neurology.* 2014;83(22):2062-2069.
18. Flanigan KM, Dunn DM, von Niederhausern A, et al. Mutational spectrum of DMD mutations in dystrophinopathy patients: application of modern diagnostic techniques to a large cohort. *Hum Mutat.* 2009;30(12):1657-1666.
19. Dubowitz V. *Muscle Biopsy: A Practical Approach.* 2nd ed. Eastbourne, England: Bailliere Tindall; 1985.
20. Beekman C, Janson AA, Baghat A, van Deutekom JC, Datson NA. Use of capillary Western immunoassay (Wes) for quantification of dystrophin levels in skeletal muscle of healthy controls and individuals with Becker and Duchenne muscular dystrophy. *PLoS ONE.* 2018;13(4):e0195850.
21. Koeks Z, Janson AA, Beekman C, et al. Low dystrophin variability between muscles and stable expression over time in Becker muscular dystrophy using capillary Western immunoassay. *Sci Rep.* 2021;11(1):5952.
22. Moen E, Bannon D, Kudo T, Graf W, Covert M, van Valen D. Deep learning for cellular image analysis. *Nat Methods.* 2019;16(12):1233-1246.
23. Weiss RB, Vieland VJ, Dunn DM, Kaminoh Y, Flanigan KM. Long-range genomic regulators of THBS1 and LTBP4 modify disease severity in duchenne muscular dystrophy. *Ann Neurol.* 2018;84(2):234-245.
24. Flanigan KM, Ceco E, Lamar KM, et al. LTBP4 genotype predicts age of ambulatory loss in Duchenne muscular dystrophy. *Ann Neurol.* 2013;73(4):481-488.
25. Bello L, Piva L, Barp A, et al. Importance of SPP1 genotype as a covariate in clinical trials in Duchenne muscular dystrophy. *Neurology.* 2012;79(2):159-162.
26. Bello L, Pegoraro E. The "Usual Suspects": genes for inflammation, fibrosis, regeneration, and muscle strength modify duchenne muscular dystrophy. *J Clin Med* 2019;8(5).
27. Simmons TR, Vetter TA, Huang N, Vulin-Chaffiol A, Wein N, Flanigan KM. Pre-clinical dose-escalation studies establish a therapeutic range for U7snRNA-mediated DMD exon 2 skipping. *Mol Ther Methods Clin Dev.* 2021;21:325-340.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Vetter TA, Nicolau S, Bradley AJ, Frair EC, Flanigan KM. Automated immunofluorescence analysis for sensitive and precise dystrophin quantification in muscle biopsies. *Neuropathol Appl Neurobiol.* 2022;48(3):e12785. doi:10.1111/nan.12785