Defining Standards for Dystrophin Quantification in DMD Muscle Biopsies

John Babiak, PhD, PTC Therapeutics

Financial Disclosure

- John Babiak – employee at PTC Therapeutics, Inc

Questions for Today

- What is the current status of assay validation and methodology for quantification of dystrophin in the context of drug development ?
  - Current methods for quantification of dystrophin have significant limitations in the ability to measure small differences in the dystrophin content of DMD muscle samples
  - The quality of biopsy samples due to sampling, handling and the extent of disease progression can negatively impact the interpretation of muscle biopsy dystrophin expression data in large clinical studies
- What are the key opportunities and knowledge gaps in the field ?

Questions for today:

How much dystrophin is in a DMD muscle ?

- Western blot: 0 – 5% of “healthy control”
- Immunostaining: 5 – 20% of “healthy control”

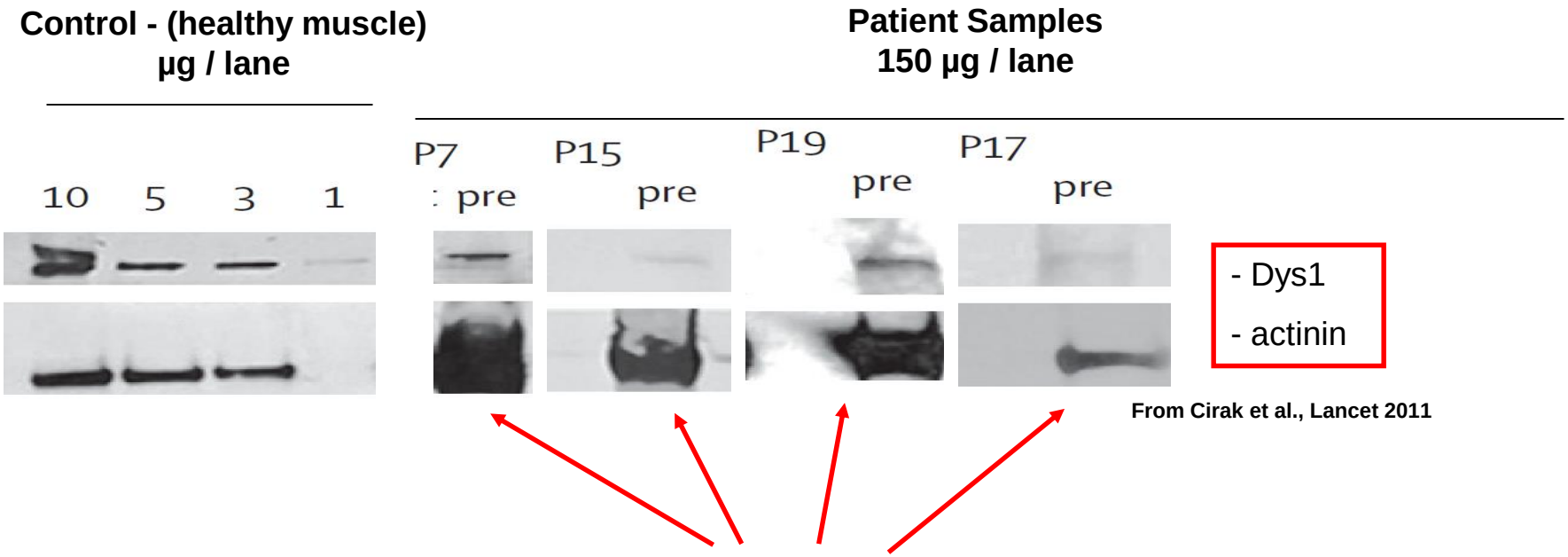
Why is there such a large difference between methods ?

- Assumptions regarding linearity of measured response
- Use of relevant reference standards (or not)
- Subtracting background or other interfering signal (or not)

Fundamental goals: Can an assay (samples + quantification method) consistently detect a small difference between two samples ?

Can you scale the assay to run a large number (100s) of samples ?

Use of a standard curve in Western Blot supports the concept that DMD muscle dystrophin is in the range of 0 – 5%



From Cirak et al., Lancet 2011

- Normalization to an internal reference muscle protein is a concern
  - A common practice is to normalize to a “loading control” muscle protein
  - This can severely impact reliability of results due to non-linearity of staining intensity of “overloaded” loading control
- Overloading of protein from DMD biopsy samples can introduce migration, transfer and staining artifacts

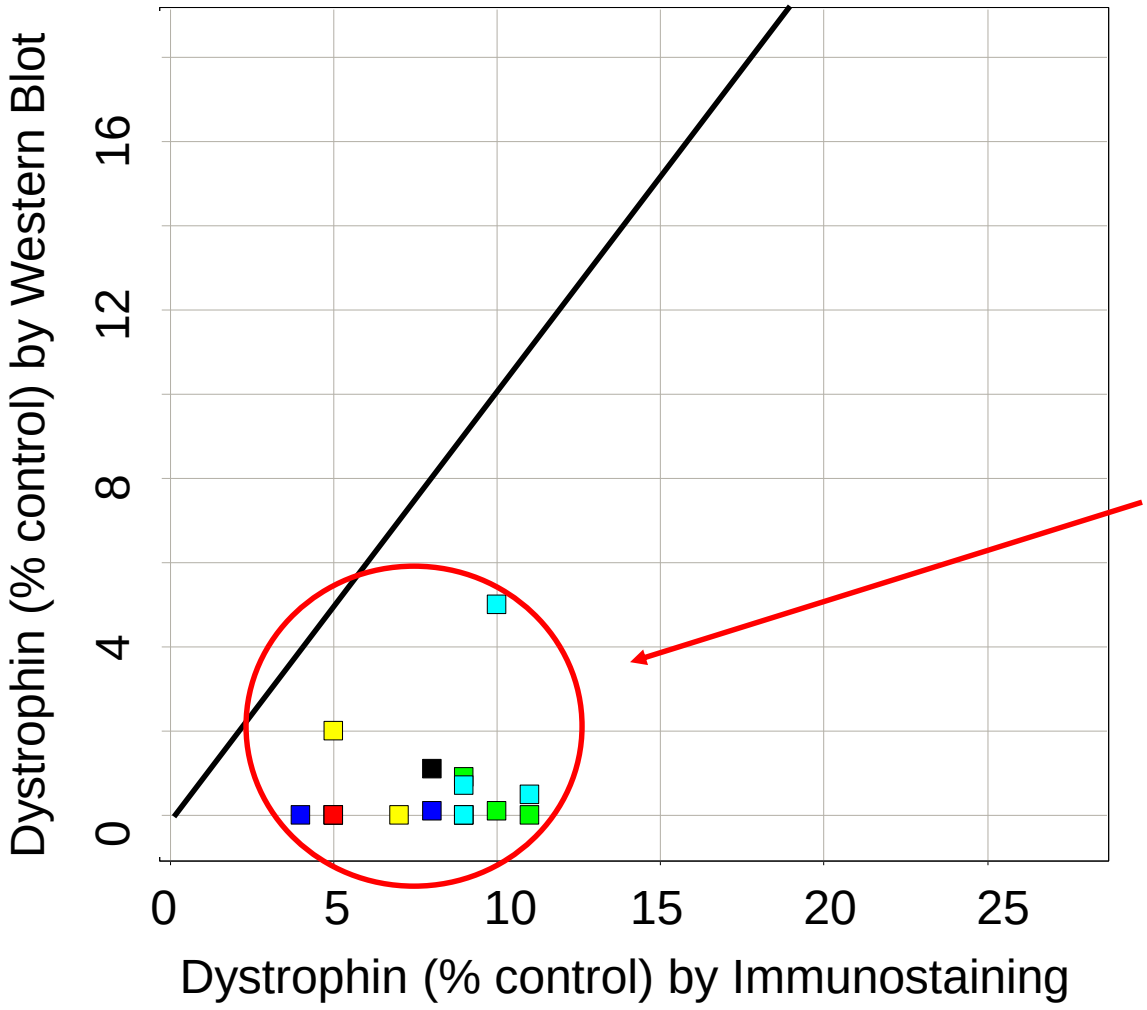
Challenges with Western blot (especially for DMD)

- Efficiency extracting a large protein from muscle tissue
- Electrophoresis and electroblot of a large protein
- Potential for degradation of dystrophin
- Need to “overload” gel to visualize dystrophin can create distortions in migration, transfer and staining, especially of “loading control” protein

What can help ?

- When available, use full-length recombinant dystrophin protein standard
  - Can run a standard curve
  - Can spike biopsies to determine extraction/detection efficiency
  - Potential to provide absolute determination of dystrophin content
- Spike biopsies with another detectable protein can also provide a useful “processing and loading control”
- Avoid normalization to an overloaded reference muscle protein

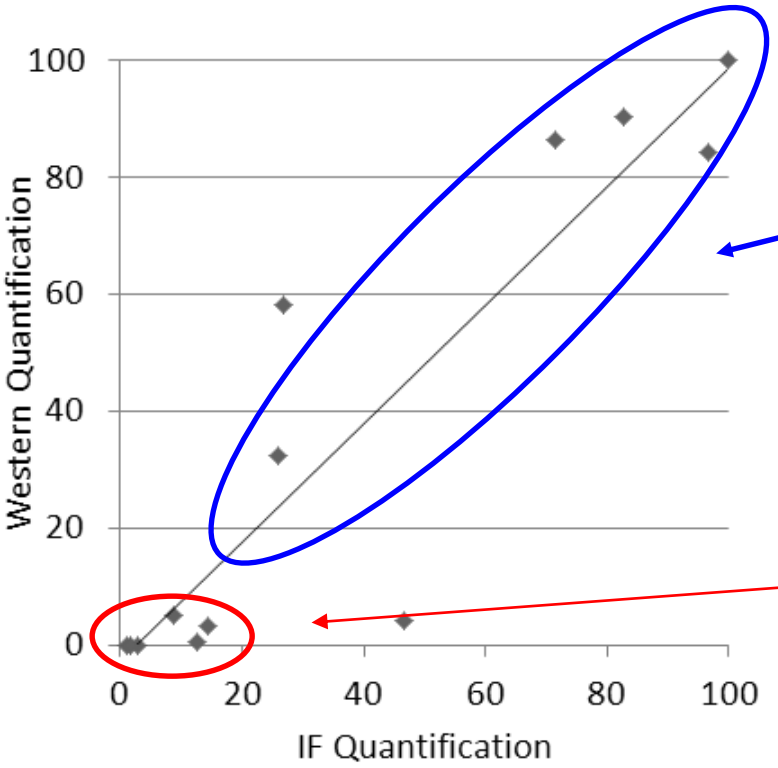
Immunostaining and Western Blot results do not correlate



Within the range relevant to DMD, immunostaining overestimates the amount of dystrophin in a muscle biopsy relative to Western blot analysis and does not correlate with WB

Adapted from data in: Cirak et al., Lancet, 2011

Forcing a correlation across a broad dynamic range reduces reliability within the range needed for DMD



Good correlation over a broad dynamic range

Very different slope over range relevant to DMD

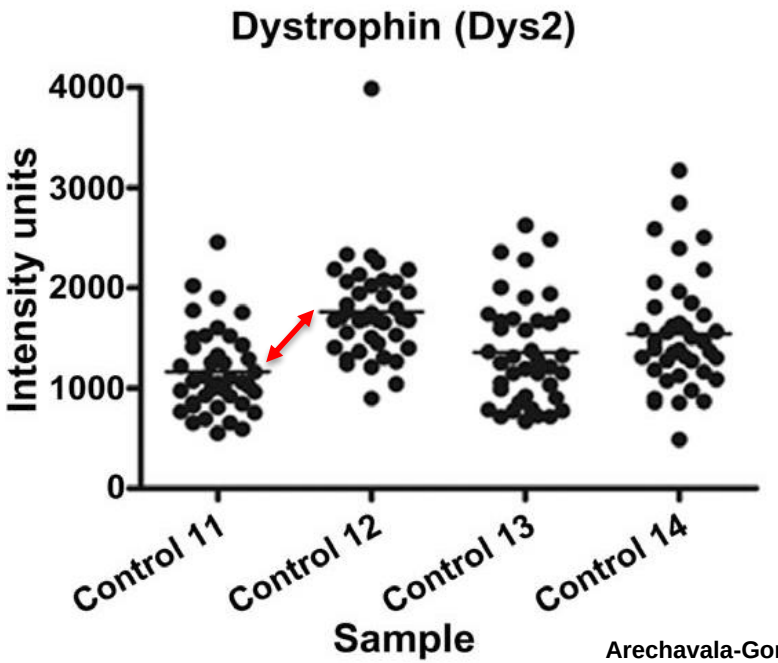
Taylor et al, Neuropath and Applied Neurobiol, 2012

Extrapolation of fit to include healthy controls distorts the linearity of response among DMD sample

The world is **not** linear, but locally it can be pretty close

Forcing long-range linearity reduces local accuracy
Forcing "healthy control" to fit on the line reduces sensitivity in measuring dystrophin in DMD samples

Normalization to “control”: There can be large variability **within** and **between** “control” biopsies

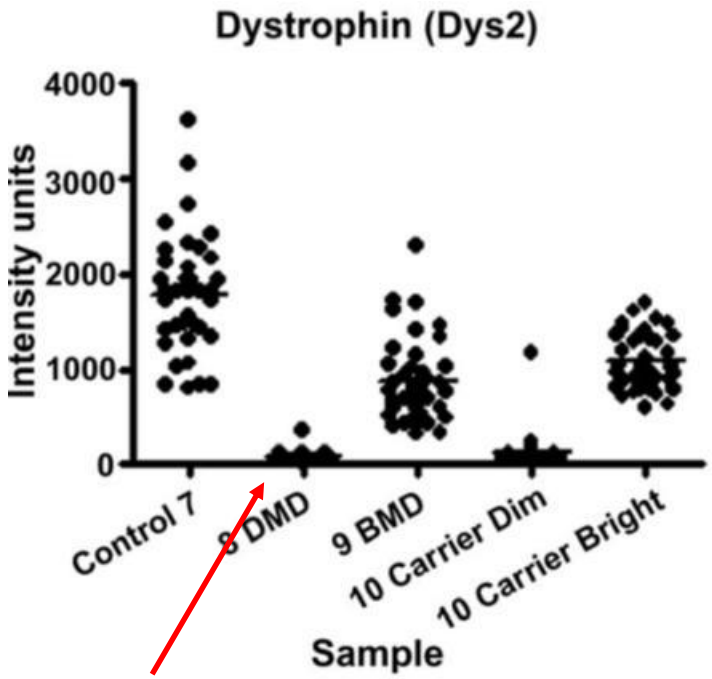


Arechavala-Gomeza et al, Neuropath and App Neurobio, 2010

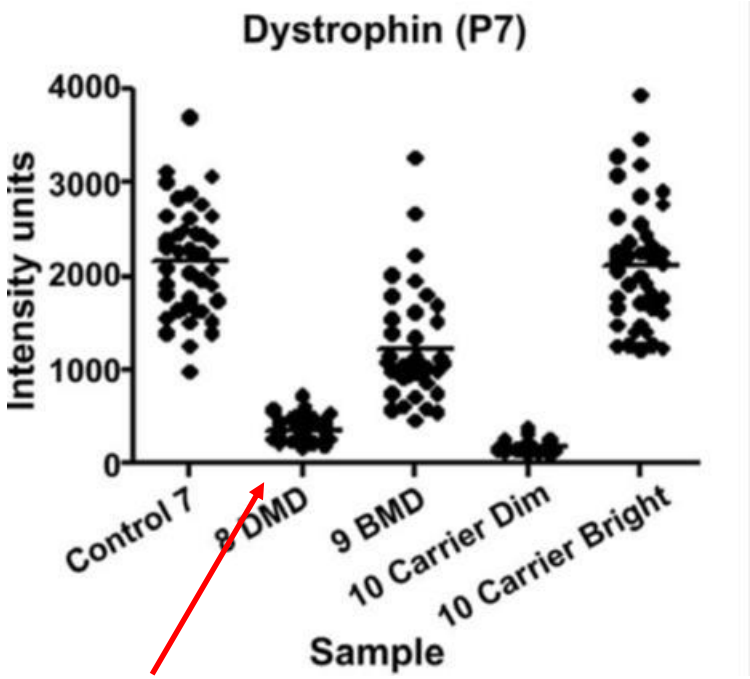
Calibration of appropriate “controls” could reduce variability across experiments

Within-sample variability and staining techniques can impact results (immunofluorescence)

- Intensity of dystrophin staining varies considerably within samples
- Different antibodies yielded 3-fold difference in dystrophin level



5% of control

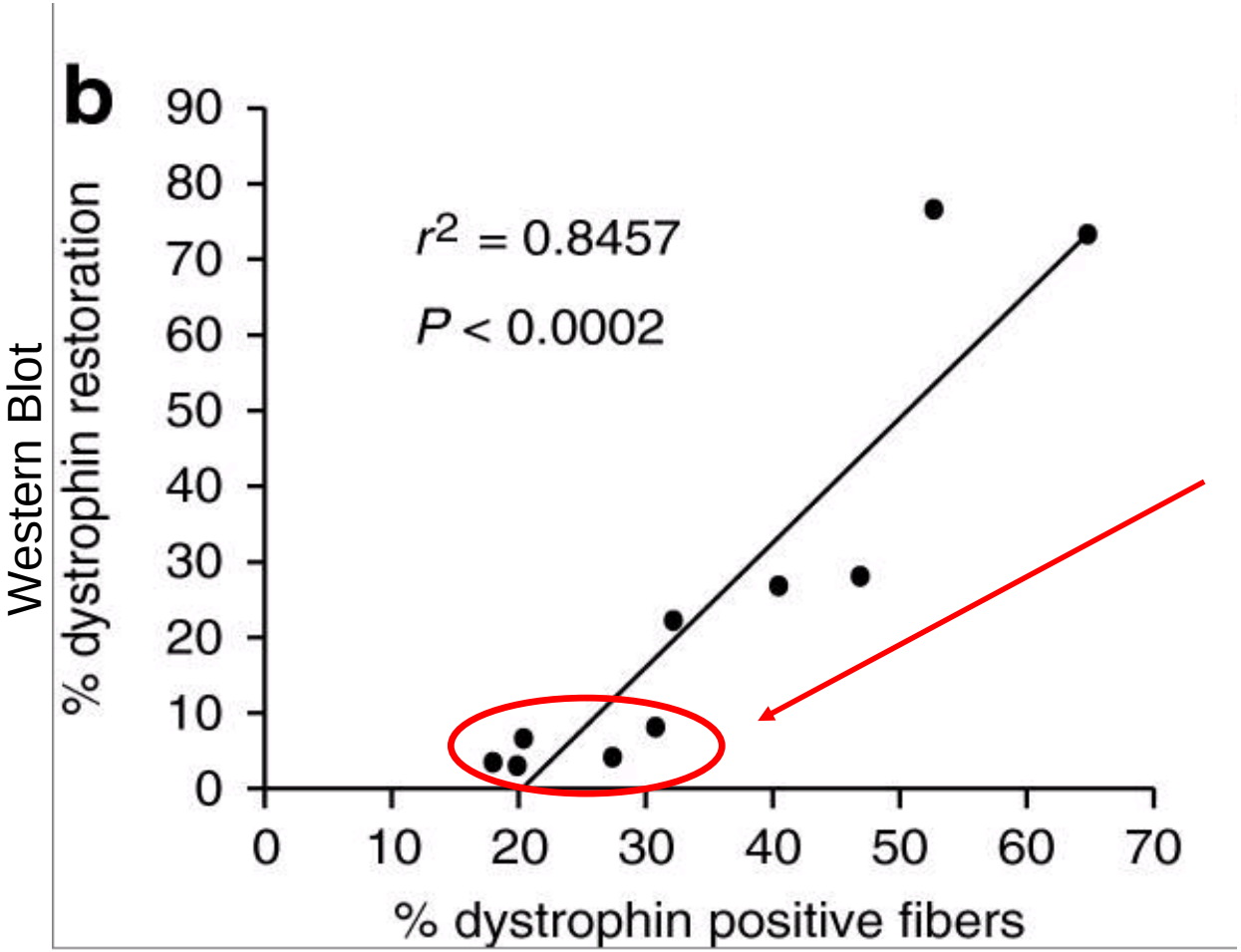


15% of control

Arechavala-Gomeza et al, Neuropath and App Neurobio, 2010

Highlights the need for reference standards in DMD range

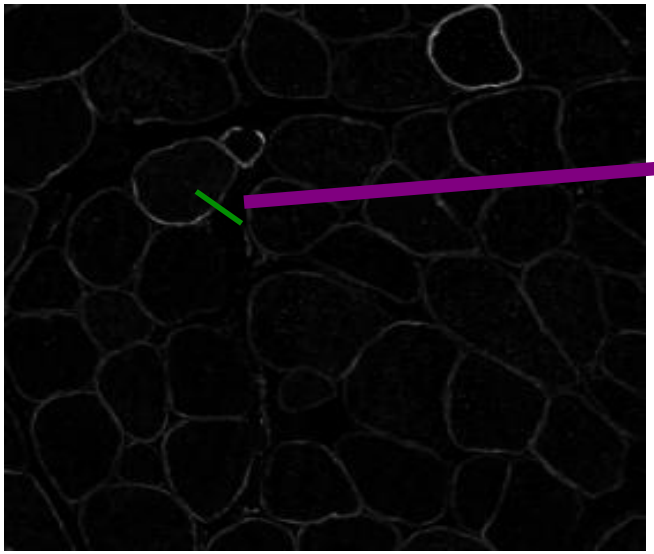
Dystrophin-positive fiber counting overestimates dystrophin expression (by Western blot) in TA muscles of PMO-treated *mdx* mice

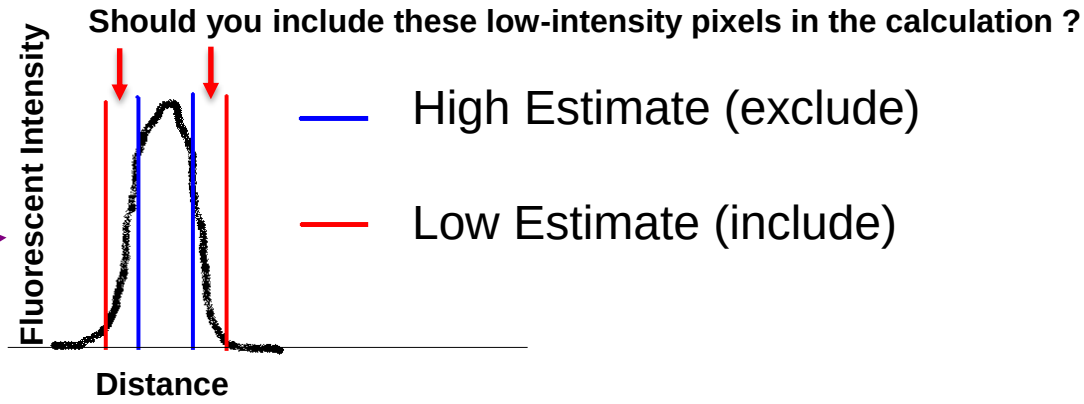


Within the range relevant to DMD, Dys+ fiber counting overestimates the amount of dystrophin in a muscle biopsy and does not correlate with Western blot analysis

Sharp et al, Molecular Therapy, 2011

Quantifying the **AVERAGE** Intensity of an IF image introduces a high degree of imprecision because of the vague definition of an **"edge"**



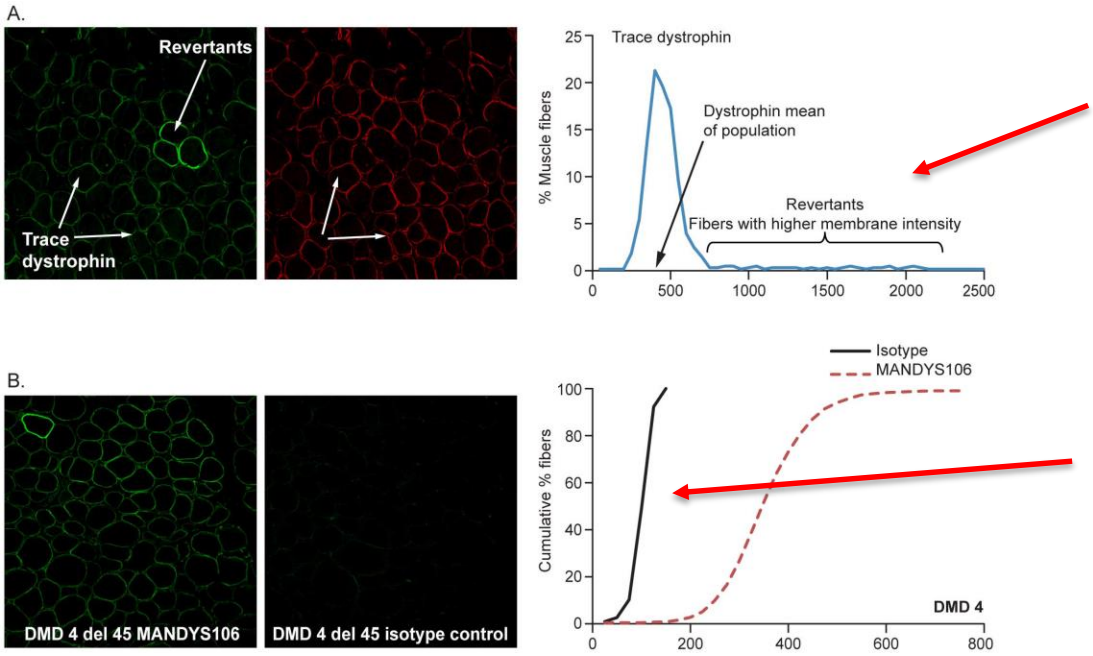


High and low estimates differ by ~50% because a large number of low-intensity pixels flank the sides of the membrane

Possible ways to reduce the impact of this limitation:

- Do not determine edge manually
- Measure TOTAL intensity of fibers of interest
- Normalize to length of membrane, not area (length x width)
- Examine pixel intensity distribution to assess extent of variability

It is possible to identify and reduce the impact of artifacts in IF analysis



Beekman et al, PLoS ONE, 2014

Identify and quantify revertant fibers

Assess background or other interfering signal intensity with isotype control antibody

Opportunities to reduce impact of background intensity

- Use isotype control to measure endogenous membrane fluorescence
- Explore use of calibrated fluorescent beads or spotting fluorescent standards onto sample to confirm linearity of fluorescent reading

Challenges with Immunostaining (especially for DMD)

- Interference from background staining and endogenous fluorescence
- Non-linearity of staining response between DMD and healthy control
- Image analysis can introduce variability (uncertainty of edge detection)
- Limited batch size of current analysis methods increases variability
- Limitations in quality of DMD samples (due to disease or sample handling)

What can help ?

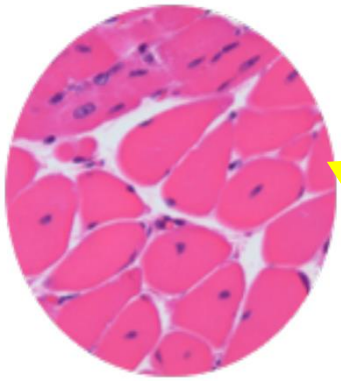
- Subtract background signal, as appropriate
- Run isotype control to measure background membrane fluorescence
- Utilize calibrated reference standard within DMD range
- Evaluate use of other calibrators (e.g. beads or spotted fluors)
- Use C-terminal-directed antibodies that detect full-length dystrophin
- Do not normalize to other muscle protein in DMD sample
- Work with gain settings appropriate to DMD, not “control” dystrophin signal
- Review algorithm for edge detection and assess potential to introduce variability
- Do not scale results to healthy control

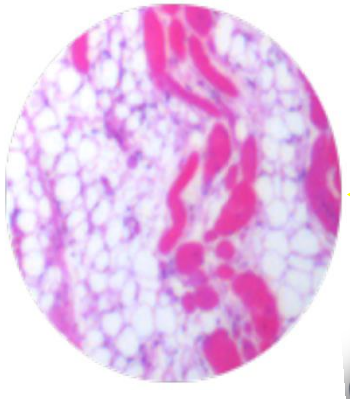
Biopsy sample integrity can confound the measurement of dystrophin in DMD muscle biopsies

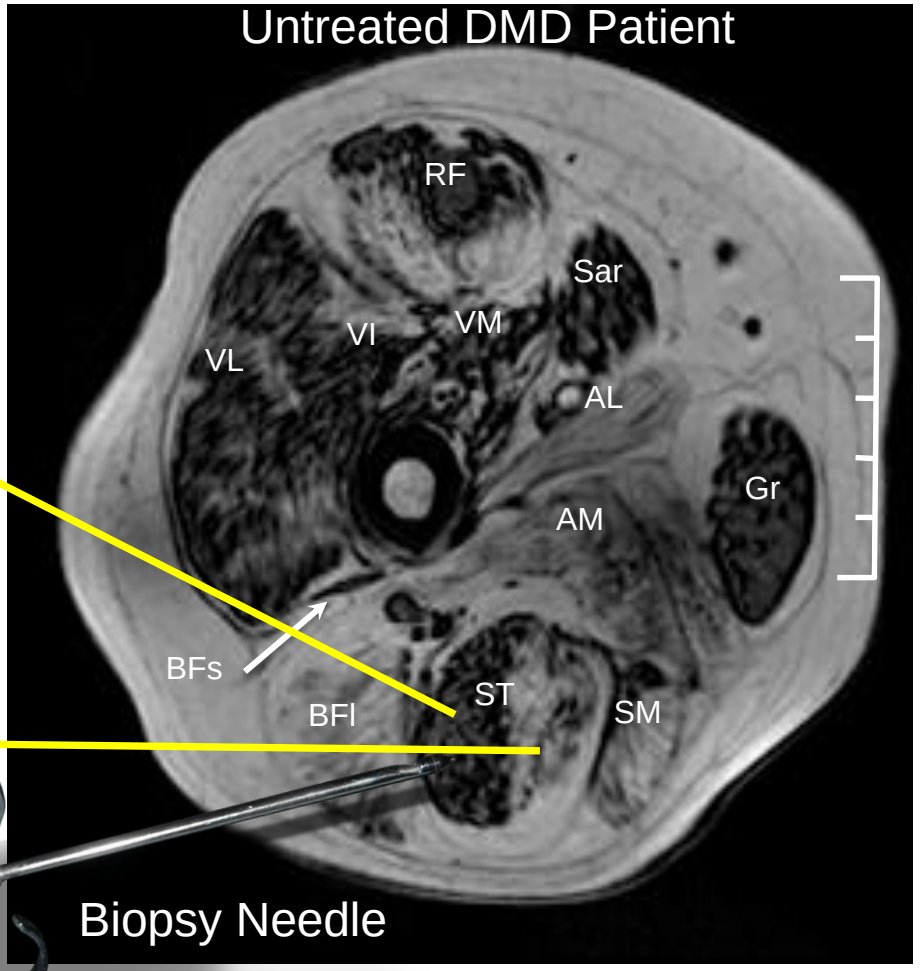
- Differences in dystrophin content between muscle groups
- Variability within a DMD muscle
- Differences in dystrophin content among fibers
- Impact of disease progression on immunostaining results
- Integrity of stored and shipped sample biopsies

Heterogeneity in Loss of Muscle Contractile Area in DMD Muscles Complicates Biopsy Interpretation

Biopsy is Invasive

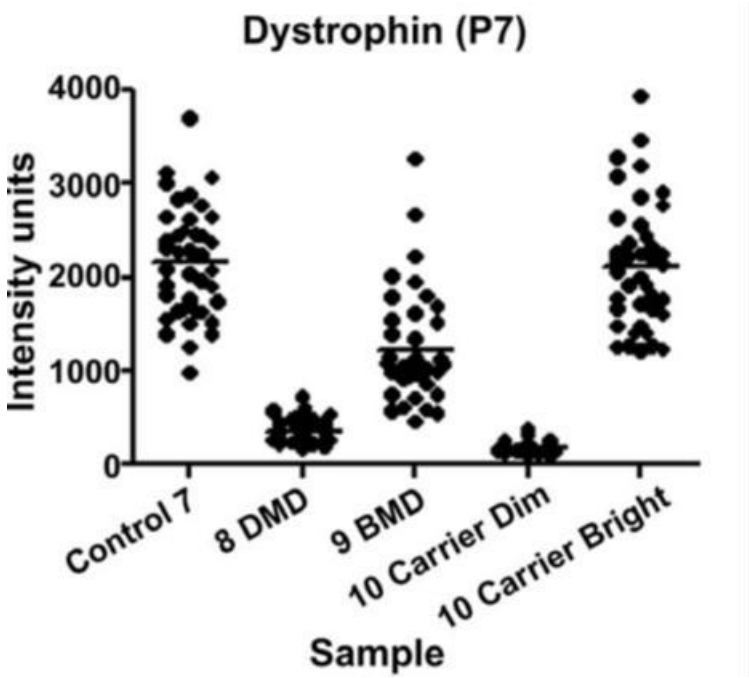




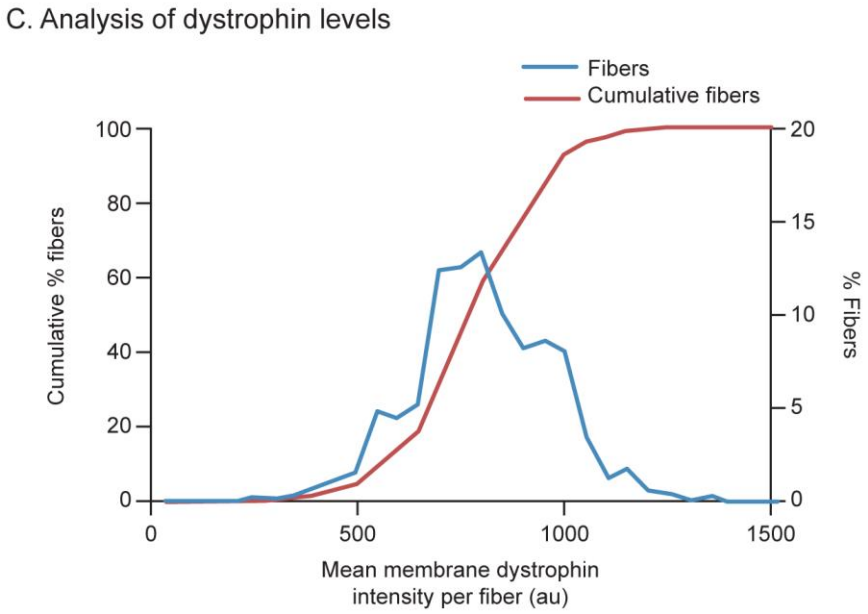


Should you biopsy "representative" or "most normal" section ?

Heterogeneity of dystrophin staining among muscle fibers is well documented

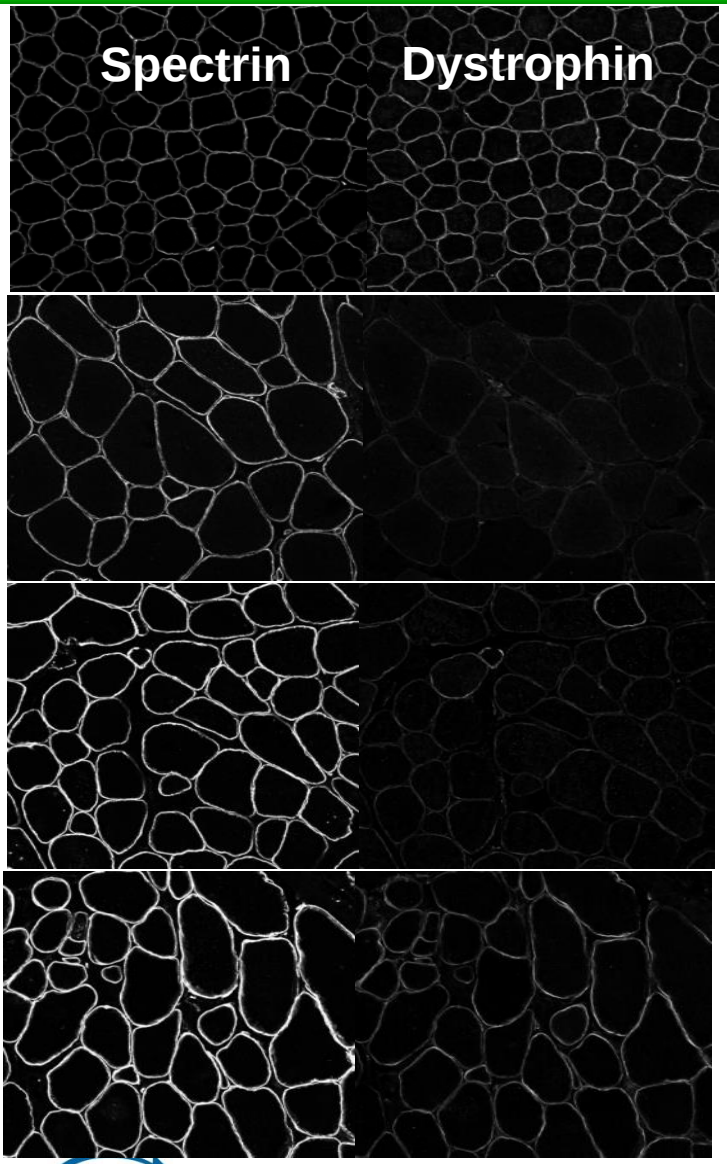


Arechavala-Gomeza et al, Neuropath and App Neurobio, 2010



Beekman et al, PLoS ONE, 2014

Dystrophin is not predictive of 6MWD



Patient A **6MWD = 474 m** **Age 10**

Patient B **6MWD = 475 m** **Age 15**

Patient C **6MWD = 292 m** **Age 12**

Patient D **6MWD = 276 m** **Age 15**

Process challenges for determining dystrophin levels in a large, multi-site clinical study

Patient

Muscle Sample Harvested

What muscle to biopsy ?
High variability in size of sample
Significant heterogeneity across entire muscle
Should you biopsy “representative” or “most normal” section ?
Invasive procedure performed multiple times in a study

Biopsy Sample Preparation at Site

Sample processing and shipping can lead to poor sample orientation and freezing artifacts

Staining/Imaging of Muscle Sample

Quantification of Dystrophin

Challenges in staining for low levels of dystrophin expression

Dystrophin quantification using dystrophic muscle adds complexity

Limited batch size of current methods

Tissue pathology and defects can introduce aberrant staining artifacts (observations in 2 clinical studies)

- "High fat and fibrotic tissue content and low muscle content or damage related to biopsy handling or shipping can prohibit immunofluorescence analysis" – Lourbakos, A., et al, WMS poster 2013

- "…the majority of muscle biopsy samples were compromised…, which confounded interpretation of the dystrophin expression results." – Bushby et al, Muscle and Nerve, 2014.

  - Freezing artifacts, n (%)
    - None = 74/342 (21.6)
    - Mild = 123/342 (36.0)
    - Severe = 145/342 (42.4)
  - Orientation, n (%)
    - Cross = 206/342 (60.2)
    - Mixed = 135/342 (39.5)
    - Longitudinal = 1/342 (0.3)
  - Endomysial fibrosis with or without fatty replacement, n (%)
    - Mild = 172/342 (50.3)
    - Moderate = 122/342 (35.7)
    - Severe = 48/342 (14.0)

Questions for Today

- What is the current status of assay validation and methodology for quantification of dystrophin in the context of drug development ?
  - Current methods for quantification of dystrophin have significant limitations in the ability to measure small differences in the dystrophin content of DMD muscle samples
  - The quality of biopsy samples due to sampling, handling and the extent of disease progression can negatively impact the interpretation of muscle biopsy dystrophin expression data in large clinical studies
- What are the key opportunities and knowledge gaps in the field ?

Conclusions – Methods and Sampling

- Potential good practices for Western blot:
  - Use a relevant standard curve
  - Avoid normalization to overloaded muscle proteins
  - Spike biopsies with detectable “processing and loading control”

- Potential good practices for Immunostaining:
  - Subtract background due to endogenous membrane fluorescence and non-specific antibody binding
  - Utilize calibrated reference standards within DMD range, NOT “healthy control”
  - Evaluate use of possible calibrators to determine linearity within relevant range
  - Utilize antibodies and fluorescence gain settings appropriate to DMD
  - Review algorithms used to detect edges and assess possible impact on results
  - Avoid normalization to other protein in DMD sample
  - Avoid fitting data to include “healthy control” results

- The quality of biopsy samples due to **sampling**, **handling** and the extent of **disease progression** can negatively impact the interpretation of muscle biopsy dystrophin expression data in large clinical studies