Dystrophin quantification
Biological and translational research implications

ABSTRACT
Objective: We formed a multi-institution collaboration in order to compare dystrophin quantification methods, reach a consensus on the most reliable method, and report its biological significance in the context of clinical trials.

Methods: Five laboratories with expertise in dystrophin quantification performed a data-driven comparative analysis of a single reference set of normal and dystrophinopathy muscle biopsies using quantitative immunohistochemistry and Western blotting. We developed standardized protocols and assessed inter- and intralaboratory variability over a wide range of dystrophin expression levels.

Results: Results from the different laboratories were highly concordant with minimal inter- and intralaboratory variability, particularly with quantitative immunohistochemistry. There was a good level of agreement between data generated by immunohistochemistry and Western blotting, although immunohistochemistry was more sensitive. Furthermore, mean dystrophin levels determined by alternative quantitative immunohistochemistry methods were highly comparable.

Conclusions: Considering the biological function of dystrophin at the sarcolemma, our data indicate that the combined use of quantitative immunohistochemistry and Western blotting are reliable biochemical outcome measures for Duchenne muscular dystrophy clinical trials, and that standardized protocols can be comparable between competent laboratories. The methodology validated in our study will facilitate the development of experimental therapies focused on dystrophin production and their regulatory approval.

Dystrophin quantification is an essential biochemical outcome measure for these trials. However, the absence of a reference standard, the large size and low expression of the protein, combined with preexisting dystrophin-positive revertant fibers and residual trace dystrophin, make accurate quantification challenging, especially when the amount of restored dystrophin is small.

Duchenne muscular dystrophy (DMD) is an X-linked neuromuscular disorder caused by mutations in the DMD gene, which prevent the expression of its product, dystrophin. The milder Becker muscular dystrophy (BMD) is also caused by DMD mutations that result in variable expression of a shorter dystrophin. Therapeutic interventions aimed at restoring dystrophin expression are in clinical trials. Dystrophin quantification is an essential biochemical outcome measure for these trials. However, the absence of a reference standard, the large size and low expression of the protein, combined with preexisting dystrophin-positive revertant fibers and residual trace dystrophin, make accurate quantification challenging, especially when the amount of restored dystrophin is small. Because regulatory authorities previously indicated that lack of consensus on the standardized methodology was an obstacle to the advancement of the field, a group of laboratories from academia and industry formed a biochemical outcome

GLOSSARY

BMD = Becker muscular dystrophy; BOM-SG = biochemical outcome measures study group; CV = coefficient of variation; DMD = Duchenne muscular dystrophy; IgG = immunoglobulin G; PBS = phosphate-buffered saline.


Correspondence to Prof. Muntoni: f.muntoni@ucl.ac.uk

Supplemental data at Neurology.org

© 2014 American Academy of Neurology. Unauthorized reproduction of this article is prohibited.
measures study group (BOM-SG) to provide a data-driven reproducible methodology for dystrophin quantification. In a pilot study comparing the sensitivity and reliability of the preferred individual laboratories’ methodologies, we found significant levels of inter- and intralaboratory variability (data not shown). Herein, we present a controlled analysis of proposed standard operating procedures for quantitative immunohistochemistry and Western blotting for evaluation of dystrophin expression. We discuss the biological significance of our data in the context of dystrophic muscle pathology. We demonstrate that data from different laboratories can be comparable, thus validating immunohistochemistry and Western blotting as biochemical biomarkers for DMD clinical trials.

METHODS Five laboratories of the BOM-SG (The Dubowitz Neuromuscular Centre, UCL, Institute of Child Health, London, UK; the Flanigan laboratory at the Center for Gene Therapy, Nationwide Children’s Hospital, Columbus, OH; Institute of Genetic Medicine, Newcastle University, UK; Institut de Myologie, UPMC UM76, INSERM U 794, CNRS UMR 7215, Paris, France; and Prosensa Therapeutics, Leiden, the Netherlands) performed blinded analysis of the same set of muscle biopsies (control [n = 2], DMD [n = 3], and BMD [n = 3]) (table 1) using standardized immunohistochemistry and Western blotting protocols.

Standard protocol approvals, registrations, and patient consents. We obtained written informed consent for the use of archived muscle tissues from all patients or guardians (as appropriate) under a protocol approved by the Nationwide Children’s Hospital institutional review board. The studies at Great Ormond Street Hospital were performed under approval by the National Research Ethics Committee (05/MRE12/32).

Muscle biopsies. We selected muscle biopsies previously archived as part of the United Dystrophinopathy Project in the Flanigan laboratory. All biopsies had been assessed for dystrophin content on a clinical or research basis and were dispensed labeled only with a blinded code, maintained in the Flanigan laboratory. Each laboratory received the same number of unfixed frozen serial 10-μm transverse muscle sections on microscope slides and an Eppendorf tube containing forty 10-μm sections of frozen muscle tissue. All laboratories were informed of the identity of the control biopsies.

Immunohistochemistry. The staining protocol, based on that of Taylor et al., was as follows:

- Transverse sections were air-dried at room temperature for 20 to 30 minutes and circled with a hydrophobic peroxidase-antiperoxidase pen.
- Primary dystrophin (rabbit C-terminal ab15277; Abcam, Cambridge, MA) and spectrin (monoclonal NCL-SPEC1; Leica Microsystems Inc., Buffalo Grove, IL) antibodies were diluted (1:400 and 1:100, respectively) in phosphate-buffered saline (PBS) and incubated with the sections for 1 hour at room temperature.
- Sections were washed (3×) in PBS for 3 minutes each.
- Each laboratory used secondary antibodies compatible with their microscope, e.g., Alexa Fluor 488 goat anti-mouse immunoglobulin G (IgG) (A11017; Molecular Probes, Eugene, OR) and Alexa Fluor 568 goat anti-rabbit IgG (A11036; Molecular Probes). These were diluted 1:500 in PBS and incubated for 30 minutes in the dark at room temperature.
- Sections were washed (3×) in PBS for 3 minutes.
- Slides were mounted using anti-fade mounting agent, e.g., ProLong Gold anti-fade reagent (Molecular Probes).

All laboratories measured dystrophin intensity using the Arechavala-Gomeza method, which measures the fluorescent intensity of 40 specific sarcolemmal regions of interest selected manually at random. Each of these regions of interest includes maximum and minimum intensity data points that are used in the data analysis. In parallel, 3 laboratories (1, 4, and 5) also quantified dystrophin using the Taylor method. This method makes use of the double staining with spectrin, another sarcolemmal protein whose level is unaffected in dystrophinopathy muscle, to create a mask that defines only the sarcolemmal area in each image. This mask allows the measurement of the intensity of the sarcolemmal area of the whole image. In addition, laboratory 4 used the Beekman method in which a spectrin mask is also used to select the sarcolemma of each individual fiber of the image. With this algorithm, the individual intensities of an average of 350 fibers are measured and the mean dystrophin intensity of the fiber population is calculated using Definiens

---

**Table 1** Sample ranking order by laboratory

<table>
<thead>
<tr>
<th>Sample</th>
<th>Lab 1</th>
<th>Lab 2</th>
<th>Lab 3</th>
<th>Lab 4</th>
<th>Lab 5</th>
<th>Sample</th>
<th>Lab 1</th>
<th>Lab 2</th>
<th>Lab 3</th>
<th>Lab 4</th>
<th>Lab 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>A (BMD)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>F (BMD)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>F (BMD)</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>D (BMD)</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>D (BMD)</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>A (BMD)</td>
<td>3</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>C (DMD)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>C (DMD)</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>E (DMD)</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>E (DMD)</td>
<td>5/6</td>
<td>5/6</td>
<td>5</td>
<td>5</td>
<td>5/6</td>
</tr>
<tr>
<td>B (DMD)</td>
<td>6</td>
<td>5</td>
<td>5</td>
<td>6</td>
<td>6</td>
<td>B (DMD)</td>
<td>5/6</td>
<td>5/6</td>
<td>6</td>
<td>6</td>
<td>5/6</td>
</tr>
</tbody>
</table>

Abbreviations: BMD = Becker muscular dystrophy; del = deletion; DMD = Duchenne muscular dystrophy; dup = duplication; ex = exon.

Samples: A = c.40_41delGA; B = dup ex 10–17; C = dup ex 7; D = del ex 3–27; E = del ex 6; F = del ex 10–44.
The key difference between these methods relates to the number of fibers measured per captured image. For a detailed protocol of each method, see e-Methods on the Neurology® Web site at Neurology.org.

Western blotting. The protocol, based on that of Taylor et al., was as follows:

- Samples were solubilized in lysis buffer (4.4 mM Tris, 9% sodium dodecyl sulfate, 4% glycerol, 5% β-mercaptoethanol).
- Loading 25 μg of protein, each laboratory used their preferred gel electrophoresis (typically 3%–8% tris-acetate gradient gels) and Western blotting equipment.
- Membranes were incubated with the anti-dystrophin primary antibody (ab15277) at 1 μg/mL overnight at 4°C in 5% milk TBS-T (TRIS buffered saline, 0.1% Tween20).
- A sarcomeric α-actinin primary antibody (Clone EA-53; Sigma, St. Louis, MO) diluted at 1:3,000 in 5% milk TBS-T was added and membranes were incubated for 1 hour at room temperature.
- Membranes were washed (3×) for 10 minutes each in PBS-T.
- Secondary antibodies compatible with the laboratories’ imaging equipment were used, e.g., horseradish peroxidase–conjugated goat anti-rabbit (1:15,000) and goat anti-mouse IgG (1:500,000) were incubated with the membranes for 30 minutes at room temperature.
- Membranes were washed (3×) for 10 minutes each in TBS-T.

Each laboratory used their preferred image acquisition equipment (e.g., Image J-based software, Odyssey infrared imaging system); the data were normalized to α-actinin and presented relative to an average of the 2 controls.

Statistical analysis. Experiments were performed in triplicate and statistical analysis was performed using GraphPad Prism version 5.03 (GraphPad Software, La Jolla, CA). The coefficient of variation (CV) was calculated using the formula CV = SD/mean × 100. For intralaboratory analysis, the CV for each laboratory for each biopsy was calculated (tables e-1 to e-6) and the CVs from the 6 biopsies were averaged for each laboratory.

The Bland-Altman plot was used to assess the agreement between different methods.

RESULTS Each laboratory ranked the samples according to the relative level of dystrophin expression determined by each technique (table 1). There was a high level of agreement among all laboratories for both immunohistochemistry and Western blotting. All laboratories identified the 3 BMD samples as having the highest dystrophin protein levels, although this top order differed between immunohistochemistry and Western blotting. By Western blot analysis, no laboratory could detect dystrophin in sample B and only 2 laboratories (3 and 4) could detect trace amounts of dystrophin protein in sample E.

Inter- and intralaboratory variability of dystrophin quantification using immunohistochemistry. From each laboratory’s data (Arechavala-Gomeza method), we calculated the mean (±SD) dystrophin levels of each sample and the CV (figure 1). Overall, the level of variability observed among the different laboratories was minimal with an average SD of 7.78 (ranging between 3.33 for sample E and 11.93 for sample A). We calculated the CV for each sample to statistically measure the degree of variation between the laboratories. A CV value of less than
20% is considered optimal. The CV values averaged 33% (ranging between 23% for sample A and 67% for sample B); samples A, C, D, and F had CV values between 20% and 30%.

We next analyzed intralaboratory variability in the same manner (figure 1). We calculated the average CV value from each laboratory (see tables e-1 to e-6 for individual data). The CV values for immunohistochemistry were below 30% for all laboratories, with laboratories 4 and 5 having low CV values of 14% and 14%, respectively.

While all laboratories were able to use the Arechavala-Gomeza method, some laboratories had access to software that enabled them to directly compare this method with additional automated methods. Three laboratories (1, 4, and 5) analyzed the same samples using the Taylor method and one laboratory (4) compared 3 different methods (Arechavala-Gomeza, Taylor, and Beekman methods). We analyzed the same images using the above intensity measurement techniques and we assessed the level of agreement between them by plotting the mean (±SD) of each sample for all techniques (figure 2A). Next, rather than calculate the correlation coefficient, which can hide a considerable lack of agreement, we plotted the data with a regression line, plotting the more automated Taylor or Beekman methods against the Arechavala-Gomeza method (figure 2B). We then selected the 2 methods used by more than one laboratory (Arechavala-Gomeza and Taylor methods), observed that the mean data from the 2 techniques were essentially identical (figure 2, A and B), and generated a Bland-Altman plot (figure 2C). This analysis shows that both methods are equivalent: the bias (the difference between the means) was only 2.103 and the 95% limits of agreement were between 10.83 and −6.63.

Inter- and intralaboratory variability of dystrophin quantification using Western blotting. We assessed the level of inter- and intralaboratory variability observed with Western blotting as above (figure 3). We observed more variability with Western blotting than immunohistochemistry with a mean SD of 15.95 (ranging between 0.89 for sample E and 33.09 for sample F). The CV values for Western blotting averaged 80% (ranging between 23% for sample F and 223% for sample E) confirming a higher degree of variability with this technique; only samples D and F had CV values nearing 20% (figure 3). The CV values were particularly affected by 2 of the samples (B and E) being at/below the limit of sensitivity; our results thus indicate that the interlaboratory variability improves as the level of dystrophin increases.

Intralaboratory variability was also more pronounced than for immunohistochemistry. Only laboratory 1 had an optimal CV value of 0.3%; laboratory 3 had the highest at 119% (figure 3).

Immunohistochemistry and Western blotting data comparison. To assess the level of agreement between the immunohistochemistry and Western blotting data, we plotted the mean (±SD) of each sample for both techniques (figure 4A), plotted the data with a regression line (figure 4B), and generated a Bland-Altman plot of the difference between the methods against their mean (figure 4C). The level of bias was −14.18 and the upper and lower limits of agreement were 10.83 and −6.63.
were 64.96 and 93.32, respectively. Although our sample size is small, the scatter of data points in figure 4C suggests that as the mean increases, the difference between the 2 methods also increases. Thus, while the immunohistochemistry and Western blotting data are somewhat comparable, the data are not in perfect agreement; this is unlikely to be attributable to technical problems but rather to the different properties of mutant dystrophin. For example, there is a large discrepancy for sample F (patient with BMD with a deletion of exons 10–44) with which the level of dystrophin quantified by Western blotting is considerably higher than that determined by immunohistochemistry (see discussion).

**DISCUSSION**

Dystrophin expression is being used as a secondary outcome measure in several clinical trials, but the lack of standardized procedures limits the ability to compare these different studies.

We set up a study in which we first standardized the methodologies for detecting dystrophin expression and then applied them to assess patients with DMD and BMD. Our data show that optimized immunohistochemistry and Western blotting were surprisingly concordant given the variable nature of dystrophinopathy biopsies (e.g., variable fibro-fatty replacement between biopsies and variable dystrophin content within serial sections of the same biopsy). We demonstrated that properly handled tissue can be distributed to multiple centers internationally to achieve comparable results. Because recent studies demonstrated that dystrophin expression can vary between different controls, we distributed control biopsies to each laboratory. In the context of clinical trials, this is a variable that will need to be considered, either by using one set of control samples, which may not be realistic, or the use of humanized mouse muscle. In any case in a clinical trial, the use of each patient’s pre- or non-treated muscle biopsy is of paramount importance because of the variable levels of revertant fibers and trace dystrophin expression in each patient.

Our study demonstrates that mean dystrophin levels obtained using 3 alternative immunohistochemical methods were comparable, suggesting that when the mean dystrophin level per biopsy is reported, the choice of which published script is used is not crucial, although extra information could be obtained with some approaches over others. We have validated the robustness, in a multicentre setting, of the Arechavala-Gomeza (5 laboratories) and Taylor (3 laboratories) methods by evaluating the precision of results generated by different equipment and operators. The Beekman method was only tested in one of the participating laboratories. One limitation of this study was that the numbers of controls and test specimens were relatively small, because it is extremely challenging to obtain human muscle biopsies of the size needed for such a comparative study.
Immunohistochemistry and Western blotting (measuring sarcolemmal and total dystrophin, respectively) give information that is not necessarily identical, but rather complementary, especially in diseased muscle. For example, in some samples (e.g., BMD sample A, c.40_41delGA), the level of dystrophin determined by both techniques was highly comparable, while with others (e.g., BMD sample F, del ex 10–44), the level of dystrophin quantified by Western blotting was significantly higher than that determined by immunohistochemistry. Considering that this patient carries a large deletion removing a significant portion of the actin-binding domain, and the ability of this mutant dystrophin to bind to the sarcolemma, we suggest that Western blotting in this case overestimates the amount of functional dystrophin. Similar findings have been reported in transgenic mds mice carrying BMD-like molecules, in which lower levels of mini-dystrophin were observed in the sarcolemmal Western blot fractions compared with mice expressing full-length dystrophin. Considering that many BMD mutations (and the equivalent DMD mutations after exon skipping) affect the 3-dimensional structure and actin-binding properties of dystrophin, capturing both the total amount of dystrophin in the homogenate as well as its localization at the sarcolemma is clearly important. Assessing dystrophin using immunohistochemistry is also important, because a different pattern of expression can lead to differences in the functional outcome irrespective of the total amount of protein present. For example, in transgenic mds mice, mice with a low, but uniform dystrophin expression have a milder phenotype than mds mice with a higher, but variable pattern.

Immunohistochemistry and Western blotting are not necessarily the most novel methods for dystrophin quantification but they remain widely available and accessible. Alternative, but less widely available techniques, such as mass spectrometry and ELISA, may be advantageous for detecting linear increments of dystrophin from very small amounts of sample. However, their use in isolation would not be desirable because of the issues related to the functionality and localization of mutant dystrophin discussed above. Based on the results of our study, we recommend that dystrophin restoration in clinical trials should be quantified using parallel techniques, which are, in hierarchy of importance: (1) sarcolemmal dystrophin quantification by immunohistochemistry, and (2) quantitative Western blotting or alternative techniques measuring total dystrophin levels in muscle homogenates such as mass spectrometry. Counting dystrophin-positive fibers is also used, but interlaboratory reliability has not been assessed; it currently relies on a qualitative rather than quantitative operational definition for "positive fibers." Nevertheless, within a single laboratory, the reproducibility of counting dystrophin-positive fibers has been indicated, although the use of a pretreatment threshold is paramount.

Our study demonstrates that when biopsy preparation and antibody protocols are standardized, multiple laboratories are able to reliably measure dystrophin expression using existing techniques. We therefore recommend the use of standardized immunohistochemical and Western blotting methods in parallel as robust biochemical outcome measures for DMD clinical trials.
AUTHOR CONTRIBUTIONS
Karen Anthony: drafting/revising the manuscript for content, study concept or design, analysis or interpretation of data, acquisition of data, statistical analysis, and study coordination. Virginia Arechavala-Gomez: drafting/revising the manuscript for content, study concept or design, analysis or interpretation of data, acquisition of data, and study coordination. Laura E. Taylor: study concept or design, contribution of vital reagents/tools/patents, and study supervision or coordination. Adeline Vuillon: analysis or interpretation of data and acquisition of data. Yunki Kaminoh: analysis or interpretation of data, acquisition of data, statistical analysis. Silvia Torelli: analysis or interpretation of data, acquisition of data, drafting/revising the manuscript for content. Lucy Feng: drafting/revising the manuscript for content, including medical writing for content, and contribution of vital reagents/tools/patents. Narinder Jangha: analysis or interpretation of data, acquisition of data. Gisèle Bonne: analysis or interpretation of data and study supervision/coordination. Maud Beuvin: analysis or interpretation of data. Rita Barresi: revising the manuscript for content, acquisition of data. Matt Henderson: revising the manuscript for content, contribution of vital reagents/tools/patents and acquisition of data. Steven Laval: analysis or interpretation of data. Afrodite Lourbakos: drafting/revising the manuscript for content, including medical writing for content, study concept or design, and analysis or interpretation of data. Gile Campion: drafting/revising the manuscript for content, including medical writing for content, analysis or interpretation of data, and study supervision or coordination. Volker Straub: study concept or design, analysis or interpretation of data, acquisition of data, study supervision or coordination, and obtaining funding. Thomas Voit: study concept and design, interpretation of data, study supervision, and obtaining funding. Caroline Sewry: drafting/revising the manuscript for content. Jennifer Morgan: drafting/revising the manuscript for content, study concept or design, analysis or interpretation of data. Kevin M. Flanigan: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data, contribution of vital reagents/tools, acquisition of data, and study supervision/coordination. Francesco Muntoni: drafting/revising the manuscript for content, including medical writing for content, study concept or design, analysis or interpretation of data, study supervision or coordination, and obtaining funding.

ACKNOWLEDGMENT
The authors gratefully acknowledge Eric Hoffman, PhD (Children’s National Medical Center, Washington, DC, additional member of the BOM-SG), Kristy Brown, PhD (Children’s National Medical Center, Washington, DC, additional member of the BOM-SG), HealthCare Gordish-Dressman, PhD (Children’s National Medical Center, Washington, DC, additional member of the BOM-SG), Alessandra Ferlini, MD, PhD (University of Ferrara, Italy, additional member of the BOM-SG), Steven Moore, MD, PhD (University of Iowa, additional member of the BOM-SG), Diane Frank, PhD (Sureptiva Therapeutics, additional member of the BOM-SG), Ellen Welch, PhD (PTC Therapeutics, additional member of the BOM-SG), Darren Chambers (University College London, technical help), and Deborah Ridout (University College London, statistical help).

STUDY FUNDING
The research leading to these results has received funding from the following grant-giving bodies: European Union’s seventh framework programme (FP7/2007-2013) under SKIP-NMD grant agreement number 201324, a PI on 3 ongoing clinical trials (PTC, PROSENSA), and receiving academic research support from EU FP 7 SKIP-NMD, SCOPE-DMD, and BIOIMAGE-NMD. C. Sewry reports royalties from 2 books—publishers Elsevier and Wiley-Blackwell; she is also a member of the editorial board of Neuromuscular Disorders and Muscle & Nerve and executive editor of Neuropathology and Applied Neurobiology. J. Morgan reports no conflict of interest. K. Flanigan has served as a site investigator for studies funded by GSK, PTC Therapeutics, Halo Therapeutics, and Prothena Therapeutics; on advisory boards for Sureptiva Therapeutics; and as a consultant to Ede Lilly. He has received research funding from the NHI, Parent Project Muscular Dystrophy, and CureDuchenne and is a member of the editorial board of Neuromuscular Disorders. F. Muntoni has served on scientific advisory boards for Roche, PTC, Summit, Italfarma, and Sureptiva Therapeutics; he serves on the scientific advisory board for Pfizer and the editorial board of Neuromuscular Disorders and Neuropathies; his institution (UCL) receives research support from the European Union, the Medical Research Council, the Wellcome Trust, the Association Française contre les Myopathies (AFM), the Muscular Dystrophy Campaign, the Great Ormond Street Hospital (GOSH) Biomedical Research Centre, GSK, Genethon, and NIH. His institutions (UCL and GOSH) receive funding for clinical trials from GSK, Trophos, Prothena, the British Heart Foundation, Summit, and has received funding for trials from AVI BioPharma and PTC Therapeutics. Go to Neurology.org for full disclosures.

Received March 14, 2014. Accepted in final form September 2, 2014.

REFERENCES


