Published in final edited form as:
Pediatr Neurol. 2015 February ; 52(2): 202–205. doi:10.1016/j.pediatrneurol.2014.09.014.

Composite biomarkers for assessing Duchenne muscular dystrophy: an initial assessment

Irina Shklyar, MD[1], **Amy Pasternak, PT**[2], **Kush Kapur, PhD**[2], **Basil T. Darras, MD**[2], and **Seward B. Rutkove, MD**[1]

[1]Departments of Neurology, Beth Israel Deaconess Medical Center, Boston, MA

[2]Boston Children's Hospital, Harvard Medical School, Boston, MA

Abstract

BACKGROUND—Compared to individual parameters, composite biomarkers may provide a more effective means for monitoring disease progression and the effects of therapy in clinical trials than single measures. In this study, we built composite biomarkers for use in Duchenne muscular dystrophy (DMD) by combining values from two objective measures of disease severity: electrical impedance myography (EIM) and quantitative ultrasound (QUS) and evaluating how well they correlated to standard functional measures.

METHODS—Utilizing data from an ongoing study of EIM and QUS in 31 DMD and 26 healthy boys aged 2–14 years, we combined data sets by first creating z-scores based on the normal subject data and then using simple mathematical operations (addition and multiplication) to create composite measures. These composite scores were then correlated to age and standard measures of function including the six-minute walk test, the North Star Ambulatory Assessment (NSAA), and handheld dynamometry.

RESULTS—Combining data sets resulted in stronger correlations with all four outcomes than for either EIM or QUS alone in six of eight instances. These improvements reached statistical significance ($p < 0.05$) in several cases. For example, the correlation coefficient for the composite measure with the NSAA was 0.79 but was only 0.66 and 0.67 (respectively) for GSL and EIM separately.

CONCLUSIONS—Arithmetically derived composite scores can provide stronger correlations to functional measures than isolated biomarkers. Longitudinal study of such composite markers in DMD clinical trials is warranted.

Correspondence to: Seward B. Rutkove, MD, Chief, Division of Neuromuscular Diseases, Department of Neurology, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, TCC-810, Boston, MA 02215, 617-667-8130(office); 617-667-3175 (fax), srutkove@bidmc.harvard.edu.

Keywords

electrical impedance myography; quantitative ultrasound; Duchenne muscular dystrophy; biomarker; outcome measure; composite

INTRODUCTION

A variety of potential therapeutic approaches are currently being studied in Duchenne muscular dystrophy (DMD), including exon-skipping strategies,[1] gene therapy,[2] myostatin inhibitors,[3] and anti-fibrotic agents.[4] Some of these, especially the exon-skipping approaches, are already demonstrating impressive potential value and may ultimately help in converting progressive DMD into a disease similar to Becker muscular dystrophy.[1] To date, potential therapies are being assessed with clinical outcome measures such as the six-minute walk test (6MWT) and the North Star Ambulatory Assessment (NSAA).[5,6] While these measures are useful, they are limited in a number of respects. First, they have inherent variability, are limited by effort and mood, and can only be completed in ambulatory boys. Moreover, such methods typically only show decline in children about 7 years of age or older and thus cannot provide data in younger children who may be most responsive to treatment; this reduces the inclusivity of most clinical trials. Moreover, these measures may not have been sensitive enough to detect therapy effects in two recent trials.[7]

Rapid, safe, and objective surrogate measures that correlate strongly to disease status could potentially find wide use in Phase II and III clinical trials in DMD. Imaging, such as MRI, have also been proposed as potential outcome measures;[8] however, it is limited by cost and lengthy image acquisition time, which may be difficult for children. Quantitative ultrasound (QUS)[9] and electrical impedance myograpy (EIM)[10] are two attractive, objective candidates for evaluating neuromuscular pathology. Ultrasound can be quantified by measuring the grayscale level (GSL), which reflects the degree of brightness in the muscle. In DMD, fibrosis and fatty infiltration result in brighter images and higher GSL values.[11] EIM is a painless, non-invasive tool that relies on the application of a small current and measurement of surface voltages. EIM detects properties of healthy muscle, including age-related increases in muscle fiber size resulting in increasing muscle capacitance, that are lost in DMD.[12] We recently studied cross-sectional data in DMD and identified that both modalities provided excellent discrimination between DMD and control subjects and correlated with the NSAA in children with DMD.[13,14] However, the two measures only correlated moderately with one another ($R_{Spearman} = -0.40$, $p=0.054$), and thus, the two methods provide complementary data on disease status. This is perhaps not unexpected since QUS relies on backscattered acoustic energy, while EIM relies on transmitted electrical energy. Accordingly, here we study the concept of creating a composite measure of disease status by combining data from these two modalities, an approach that has been used with success in magnetic resonance imaging studies in multiple sclerosis.[15,16] This strategy has the potential to result in new, sensitive outcome measures that could be used in future clinical trials to facilitate drug development in DMD.

METHODS

Subjects and recruitment

The recruitment process has been described previously.[13] Briefly, the Boston Children's Hospital Institutional Review Board approved the protocol. Patients provided written consent, and children provided verbal assent. Boys with DMD and healthy boys aged 2 to 14 were recruited.

EIM and ultrasound measurements

The methods for GSL and EIM acquisition have also been described previously.[13] Briefly, six muscles, including deltoids, biceps, wrist flexors, quadriceps, tibialis anterior, and medial gastrocnemius were measured transversely relative to the long axis of each muscle. Each subject underwent a maximum of three measurements at baseline, 6, and 12 months. EIM measurements were obtained with the Imp SFB7 (Impedimed, Inc, Sydney, Australia) using a custom hand-held array,[17] with three different probe sizes being used depending on the child's size. The array dimensions were: Small: 4 X 1.5cm; Medium 5 X 2cm; Large: 7 X 2.5 cm. US images were obtained using the Terason t3000 system (Teracorp, Inc, Burlington, MA) with a 10 MHz probe. All images were converted to JPEG files and analyzed using Matlab® (MathWorks, Inc, Natick, MA) to obtain the brightness of the region of interest, measured as median grayscale level (GSL).[18] The region of interest (ROI) was defined as a region of fixed dimensions (130 pixels × 64 pixels) and placed in the area of muscle directly below the subcutaneous fat layer. For both EIM and US, measurements were performed on the same muscles and locations. For this analysis, and for simplicity, data from all six muscles were averaged and the 6-muscle average values for EIM and QUS used in all analyses.

Standard functional measures

The 6MWT, NSAA, and handheld dynamometry (HHD) were all performed by an experienced pediatric physical therapist (AP). For HHD, shoulder abduction, elbow flexion, forearm flexion, knee extension, foot dorsiflexion and foot plantar flexion were measured each three times. The highest value obtained for each muscle was then averaged across all the muscles to provide a single average HHD score for each subject.

Data analysis, including creation of z- and composite scores

EIM phase is measured in degrees, and GSL is dimensionless. Thus, in order to create a composite score from these independent variables, we developed z-scores based on the healthy subject data, in which individual measurement values are replaced by values that are measured in standard deviations relative to the group mean. Thus a z-score of +0.5 for a boy with DMD would indicate that the value was 0.5 standard deviation above the group mean for healthy boys; a z-score of −0.5 would be 0.5 standard deviation below the healthy boy mean. To do so, we first confirmed a relatively normal distribution for the EIM and GSL healthy subject data. The 6-muscle average values for all healthy subjects and the associated standard deviations for both EIM and GSL were then obtained separately. The difference between raw individual DMD patient data points and the mean healthy subject value was

calculated and then divided by the SD for the normal subject data, providing a z-score for each DMD patient's averaged 6-muscle data point for both EIM and GSL separately.

However, worsening disease in US is accompanied by elevations in GSL and thus positive z-scores, whereas worsening of disease with EIM results in lower phase values and thus negative z-scores. Accordingly, EIM scores were multiplied by −1 to make the direction of change for both z-scores consistent. The final composite score was created via simple arithmetic combinations: either adding or multiplying the EIM and GSL z-scores. The output was then correlated to age, NSAA, 6MWT, and HHD (via Spearman analysis) and the results compared to correlations for the individual EIM and US data. Steiger's Z test was used to compare rho values to determine if the differences between values were significant at the $p < 0.05$ level.[19]

RESULTS

Subject Demographics

We obtained EIM and QUS measurements on 31 subjects with DMD and 26 healthy controls in which both sets of data were acquired. DMD subjects had a median age of 7.81 +/− 3.42 years, and healthy boys had a median age of 7.40 +/− 2.6years (t-test, p=0.89). Data was included from a total of 65 visits with the DMD patients and 64 visits with the normal subjects. There were 30 data points for NSAA measurements, 17 for 6MWT, and 14 for HHD testing (since children could only undergo age-appropriate testing).

Correlations with surrogate measures of disease

The results of a comparison between correlation analyses of GSL and EIM and the composite scores are summarized in Table 1 and Figure 1. In short, in 6 out of 8 combinations of data sets, both approaches for creating composite scores (adding and multiplying) resulted in stronger correlation coefficients as compared to the individual EIM and GSL values, although the improvements were generally modest. Nonetheless, based on the Steiger's Z scores, several composite measures did achieve statistical significance beyond the individual parameters. The improvements were most marked for the correlations with the NSAA, where correlation for the multiplication composite measure with NSAA (rho=0.79, p<0.001) was significantly higher than both of the original GSL (rho=0.656, p<0.001) and phase (rho=0.671, p<0.001) parameters (Multiplication/GSL: z=2.04, p=0.02; Multiplication/Phase: z=1.75, p=0.04). The sum correlation likewise was significantly higher for NSAA than the original GSL (z=2.11, p=0.02), but the improvement over EIM phase did not reach significance (z=0.42, p=0.34). In the case of age, the difference in rho values between GSL and the sum composite reached significance (z=1.82, p=0.03). Additionally, the relationship between the 6MWT and the multiplication composite reached significance whereas the correlations with individual EIM and GSL parameters did not. However, there was not a significant difference between the composite rho value and those of either individual parameter (p>0.19).

DISCUSSION

These results confirm the basic premise that, by combining disparate data sets, it is possible to create biomarkers that have stronger correlations to standard functional measures and age the original data in isolation. At first, it might seem counterintuitive that the combination can be greater than the sum of the two parts; however, such combinations of values have been used in other neurological disorders, including measures of disease severity in multiple sclerosis and Alzheimer's disease.[16,20,21] In addition, as shown in our previous study, EIM and GSL only correlate moderately one with another,[13] and thus, likely provide distinct information about DMD pathology. Accordingly, combining the data sets logically might yield a measure with a stronger association with functional measures than either one alone.

It is interesting that our data show the strongest correlations for EIM and GSL, both separately and as a composite measures, with the NSAA and handheld dynamometry. The reason for this is not clear, but it is possible that the NSAA captures more functional aspects than measuring distance walked. Similarly, dynamometry captures strength data from both the upper and lower extremities. Why the EIMG-GSL composite correlation with NSAA showed the greatest improvement as compared to the other functional measures is unclear.

While the concept of creating z-scores and combining data is appealing in its relative simplicity and ease of implementation, it is not the only approach to create composite values. Another potentially more powerful method is by utilizing machine learning techniques, including support vector machines, to combine different data sets.[22,23] However, such an approach is challenging to use in this analysis for several reasons. First, to perform such analyses, it is necessary to have very large data sets such that they can be split into two groups: one to train the machine and the other to test the machine. In addition, such machines are easiest to apply to basic classification problems—in other words simply discriminating healthy subjects from controls—not in improving correlations to standard markers as was performed here. Finally, such approaches would likely prove impractical for widespread implementation in multi-site clinical trials.

In this analysis, we chose to build a composite measure from one electrophysiologic and one imaging biomarker. However, it is also possible to incorporate other types of biomarkers, including functional outcome measures. So to some extent, our rationale for studying and presenting this work is not only to highlight the specific potential value of combining EIM and QUS, but rather, to describe a simple approach for combining disparate sets of data in order to create markers that are of higher relevance. Importantly, the United States Food and Drug Administration has specifically identified the concept of composite biomarkers as useful and have offered approaches for their qualification as approved biomarkers in clinical trials.[24]

Our results, while statistically significant, demonstrate only modest improvement in the strength of the correlations with functional measures. More importantly, the main value in creating these composite biomarkers will not be to create stronger correlations with known functional measures in an essentially cross-sectional fashion, as achieved here, but rather to provide better outcomes longitudinally. Thus, our ultimate goal will be to see if these

biomarkers are more effective at identifying and tracking disease progression than single biomarkers alone. This question will ultimately be addressed using data from our ongoing longitudinal study in DMD patients. One potential concern is that whereas the correlation coefficients with functional outcomes may improve, combining two data sets could potentially lead to increased noise in the longitudinal data set, since each measure has its own associated variability. Hence, it will be critical to pursue this more challenging longitudinal analysis before determining whether such composite measures are of truly valuable in a clinical trials context.

In summary, composite measures emerged as parameters with stronger correlations with age and with three functional measures than compared with either original parameter alone (EIM phase or GSL). Because there is a need for more responsive and clinically meaningful outcome measures that can be used in children over a wide range of age and disability, a composite measure such as that created from EIM and GSL is especially attractive. We believe that where possible, future trials in DMD should attempt to include EIM-GSL composite measures to test this concept further in a prospective fashion. And since all clinical longitudinal trials result in an abundance of competing data sets, we encourage other investigators to study how combinations of other obtained measures may improve upon individual parameters.

Acknowledgments

This work was funded by the NIH (R01AR060850-02). Irina Shklyar's participation was funded by the Howard Hughes Medical Institute (HHMI) and Duchenne Research Fund (DRF).

References

1. Cirak S, Feng L, Anthony K, et al. Restoration of the dystrophin-associated glycoprotein complex after exon skipping therapy in Duchenne muscular dystrophy. Mol Ther. 2012; 20(2):462–467. [PubMed: 22086232]
2. Nelson SF, Crosbie RH, Miceli MC, Spencer MJ. Emerging genetic therapies to treat Duchenne muscular dystrophy. Curr Opin Neurol. 2009; 22(5):532–538. [PubMed: 19745732]
3. Amthor H, Hoogaars WM. Interference with myostatin/ActRIIB signaling as a therapeutic strategy for Duchenne muscular dystrophy. Curr Gene Ther. 2012; 12(3):245–259. [PubMed: 22554312]
4. Zhou L, Lu H. Targeting fibrosis in Duchenne muscular dystrophy. J Neuropathol Exp Neurol. 2010; 69(8):771–776. [PubMed: 20613637]
5. McDonald CM, Henricson EK, Han JJ, et al. The 6-minute walk test as a new outcome measure in Duchenne muscular dystrophy. Muscle Nerve. 2010; 41(4):500–510. [PubMed: 19941337]
6. Mazzone ES, Messina S, Vasco G, et al. Reliability of the North Star Ambulatory Assessment in a multicentric setting. Neuromuscul Disord. 2009; 19(7):458–461. [PubMed: 19553120]
7. Hoffman EP, Connor EM. Orphan drug development in muscular dystrophy: update on two large clinical trials of dystrophin rescue therapies. Discovery medicine. 2013; 16(89):233–239. [PubMed: 24229740]
8. Finanger EL, Russman B, Forbes SC, Rooney WD, Walter GA, Vandenborne K. Use of skeletal muscle MRI in diagnosis and monitoring disease progression in Duchenne muscular dystrophy. Phys Med Rehabil Clin N Am. 2012; 23(1):1–10. [PubMed: 22239869]
9. Pillen S, van Dijk JP, Weijers G, Raijmann W, de Korte CL, Zwarts MJ. Quantitative gray-scale analysis in skeletal muscle ultrasound: a comparison study of two ultrasound devices. Muscle Nerve. 2009; 39(6):781–786. [PubMed: 19301363]

10. Rutkove SB. Electrical Impedance Myography: Background, Current State, and Future Directions. Muscle Nerve. 2009; 40:936–946. [PubMed: 19768754]
11. Jansen M, van Alfen N, Nijhuis van der Sanden MW, van Dijk JP, Pillen S, de Groot IJ. Quantitative muscle ultrasound is a promising longitudinal follow-up tool in Duchenne muscular dystrophy. Neuromuscul Disord. Apr; 2012 22(4):306–317. [PubMed: 22133654]
12. Li J, Geisbush TR, Rosen GD, Lachey J, Mulivor A, Rutkove SB. Electrical impedance myography for the *in* and *ex vivo* assessment of muscular dystrophy (mdx) mouse muscle. Muscle Nerve. 201; 49(6):829–35. [PubMed: 24752469]
13. Rutkove SB, Geisbush TR, Mijailovic A, et al. Cross-sectional evaluation of electrical impedance myography and quantitative ultrasound for the assessment of Duchenne muscular dystrophy in a clinical trial setting. Ped Neurology. 2014; 51:88–92.
14. Rutkove SB, Darras BT. Electrical impedance myography for the assessment of children with muscular dystrophy: a preliminary study. Journal of physics. Conference series. 2013; 434(1)
15. Poonawalla AH, Datta S, Juneja V, et al. Composite MRI scores improve correlation with EDSS in multiple sclerosis. Mult Scler. Sep; 2010 16(9):1117–1125. [PubMed: 20813778]
16. Cutter GR, Baier ML, Rudick RA, et al. Development of a multiple sclerosis functional composite as a clinical trial outcome measure. Brain : a journal of neurology. 1999; 122 (Pt 5):871–882. [PubMed: 10355672]
17. Narayanaswami P, Spieker AJ, Mongiovi P, Keel JC, Muzin SC, Rutkove SB. Utilizing a handheld electrode array for localized muscle impedance measurements. Muscle Nerve. 2012; 46(2):257–263. [PubMed: 22806375]
18. Zaidman CM, Wu J, Wilder S, Darras BT, Rutkove SB. Minimal training is required to reliably perform quantitative ultrasound of muscle. Muscle Nerve. 2014; 50:124–8. [PubMed: 24218288]
19. Meng XL, Rosenthal R, Rubin DB. Comparing correlation coefficients. Psych Bulletin. 1992; 111:172–175.
20. Poonawalla AH, Datta S, Juneja V, et al. Composite MRI scores improve correlation with EDSS in multiple sclerosis. Mult Scler. Sep; 2010 16(9):1117–1125. [PubMed: 20813778]
21. Faux NG, Ritchie CW, Gunn A, et al. PBT2 rapidly improves cognition in Alzheimer's Disease: additional phase II analyses. Journal of Alzheimer's disease. 2010; 20(2):509–516.
22. Rizk-Jackson A, Stoffers D, Sheldon S, et al. Evaluating imaging biomarkers for neurodegeneration in pre-symptomatic Huntington's disease using machine learning techniques. NeuroImage. 2011; 56(2):788–796. [PubMed: 20451620]
23. Srivastava T, Darras BT, Wu JS, Rutkove SB. Machine learning algorithms to classify spinal muscular atrophy subtypes. Neurology. 2012; 79(4):358–364. [PubMed: 22786588]
24. FDA. Biomarker Qualification Context of Use. http://www.fda.gov/drugs/developmentapprovalprocess/drugdevelopmenttoolsqualificationprogram/ucm284620.htm

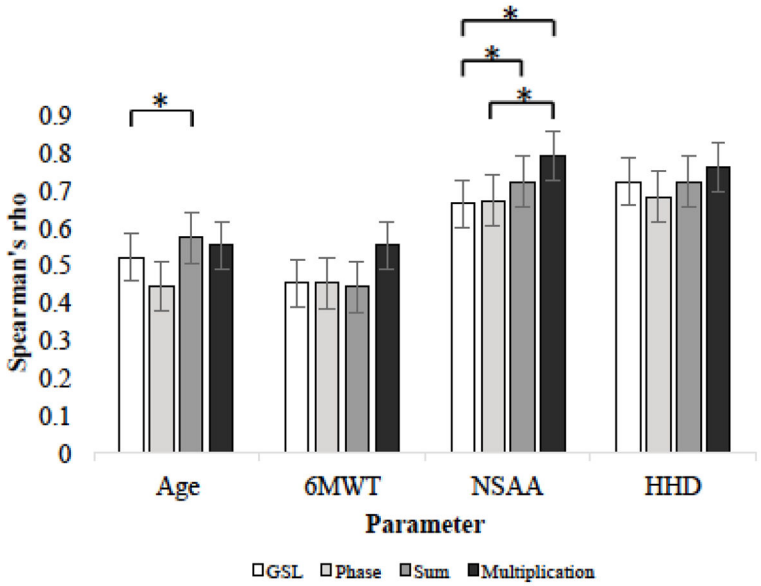

Figure 1.

Column plot summarizing the correlation coefficients for the single and composite measures, including significant differences in correlation coefficients (using Steiger's Z test). * $p < 0.05$

Table 1

Spearman rho values reflecting correlations between GSL, EIM phase, and EIM-GSL composites with age and functional measures in the DMD patients alone.

DMD surrogates of severity	Gray-scale Level	Phase 50 kHz	Sum	Multiplication
Age	0.52***	0.44***	0.57***	0.55***
Six-minute walk test	−0.45	−0.45	−0.44	−0.55*
North Star Ambulatory Assessment	−0.66***	−0.67***	−0.72***	−0.79***
Handheld dynamometry	−0.72**	−0.68**	−0.72**	−0.76**

*<0.05,

**<0.01,

***<0.001