Title: Disease duration and disability in dysferlinopathy can be described by muscle imaging using heatmaps and random forests.

Authors:
David Gómez-Andrés MD, PhD[1]; Jorge Díaz MD[2]; Francina Munell MD, PhD[1]; Ángel Sánchez-Montáñez MD[3]; Irene Pulido-Valdeolivas MD, PhD[4]; Lionel Suazo MD[2], Cristián Garrido MT[2]; Susana Quijano-Roy MD, PhD [5]; Jorge A. Bevilacqua MD, PhD[6]

Affiliations:

1. Child Neurology Unit. National Reference Unit for Rare Neuromuscular Disorders. Vall d'Hebron University Hospital. VHIR Vall d'Hebron. Research Unit. ERN EURO-NMD. ERN-RND. Barcelona. Spain.
2. Medical Imaging Center. University of Chile Clinical Hospital (HCUCH). Santiago. Chile.
3. Pediatric Neuroradiology Unit. National Reference Unit for Rare Neuromuscular Disorders. Vall d'Hebron University Hospital. VHIR Vall d'Hebron. Research Unit. ERN EURO-NMD. ERN-RND. Barcelona. Spain.
4. Visual Pathway Laboratory. Neuroimmunology Center. Neurology Department. Biomedical Research Center August Pi i Sunyer (IDIBAPS). Hospital Clínic Barcelona. University of Barcelona. Barcelona. Spain. Laboratorio de la Vía Visual. Centro de Neuroinmunología y servicio de Neurología Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS). Hospital de Barcelona. Universitat de Barcelona. Barcelona, España.
5. APHP- Neurology and Intensive Care Department. University Hospital Raymond Poincaré, Garches; U1179 Versailles University (UVSQ); Neuromuscular Disorders Reference Center of Nord-Est-Île de France (FILNEMUS), ERN Neuro-NMD. France.
6. Neuromuscular Unit. Department of Neurology and Neurosurgery. University of Chile Clinical Hospital (HCUCH) and Department of Anatomy and Legal Medicine. Faculty of Medicine, University of Chile. Santiago. Chile.

Co-corresponding authors:
David Gómez Andrés: david_gomez@vhebron.net
Jorge Bevilacqua: jbevilac@med.uchile.cl

Acknowledgements: This project was supported by grant FONDECYT 1151383 to JAB and JD. DGA and SQR were members of WG1 of Cost Action BM1304, "Applications of MR imaging and spectroscopy techniques in neuromuscular disease: collaboration on outcome measures and pattern recognition for diagnostics and therapy development. FM, ASM and JB were also participants in this Cost Action. We thank Dr. Yamile Corredoira for reviewing the English of the manuscript.

Running title: Dysferlin MRI, disability & ML

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/mus.26403

Ethical statement: We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with such guidelines

Disclosure of conflict of interest: None of the authors has any conflict of interest to disclose.

Abstract (149 words):

INTRODUCTION: The manner in which imaging patterns change over the disease course and with increasing disability in dysferlinopathy is not fully understood.

METHODS: Fibroadipose infiltration of 61 muscles was scored based on whole-body MRI of 33 patients with dysferlinopathy and represented in a heatmap. We trained random forests to predict disease duration, motor function measure dimension 1 (MFM-D1) and modified Rankin scale (mRS) based on muscle scoring and selected the most important muscle for predictions.

RESULTS: Heatmap delineates positive and negative fingerprints in dysferlinopathy. Disease duration is related to infiltration of infraspinatus, teres major-minor and supraspinatus muscles. MFM-D1 decreases with higher infiltration of teres major-minor, triceps and sartorius. mRS is related to infiltration of vastus medialis, gracilis, infraspinatus and sartorius.

DISCUSSION: Dysferlinopathy shows a recognizable muscle MRI pattern. Fibroadipose infiltration in specific muscles of the thigh and the upper limb appears to be an important marker for disease progression.

Keywords:

Muscle imaging, random forest, machine learning, dysferlin, heatmap, disability

Introduction

Dysferlinopathies are a complex group of muscle disorders caused by mutations in the dysferlin gene (*DYSF*)[1-4]. The most common clinical phenotypes are limb-girdle muscular dystrophy 2B (LGMD2B) and Miyoshi myopathy (MM) – in which a more prominent distal phenotype occurs –, but other less frequent phenotypes have also been described[1-4]. The diagnosis of dysferlinopathy is mainly based on the identification of mutations in *DYSF* gene or the detection of dysferlin deficiency in muscle or leukocytes[5]. The role of muscle imaging in diagnosis is increasingly relevant, but some areas of uncertainty still remain to be clarified[6-15].

Dysferlinopathy is a highly heterogeneous disorder. Consequently, the expression of muscle degeneration on muscle magnetic resonance imaging (MRI) is also highly variable. However, this variability in muscle imaging is still to be fully defined[12,13], even in previous studies involving large cohorts of patients[16]. This complicates the interpretation of muscle imaging across the spectrum of disease severity, evolution and expressivity. In addition, due to such uncertainty, the design of quantitative MRI studies is also limited in dysferlinopathies. Muscle selection in these studies is currently based on global criteria that do not take into account the particularities of muscle fibroadipose infiltration in dysferlinopathies. Consequently, muscles in which changes in infiltration are not related to the long-term course of the disease are included in the quantitative MRI study design, resulting in misallocation of time and resources.. Another consequence of the latter is that some muscles that may be key markers of disease progression are not evaluated, thus undermining the potential usefulness of quantitative MRI as a biomarker in dysferlinopathies.

Heatmaps are graphical representations for pattern visualization that have been previously used to study muscle disorders. In the present work, heatmaps were used to

analyse MRI data on muscle fatty infiltration from a group of patients with genetically confirmed dysferlinopathy. Such descriptive approach aimed at defining the variable range of muscle infiltration across the spectrum of disease severity and evolution. In addition, the use of random forests[17], a machine-learning approach that enhances the reproducibility and objectivity of such selection, aimed at defining those muscles that are most associated with disease duration and disability.

Methods

Patients

Between 2011 and 2017, 33 patients with dysferlinopathy were evaluated clinically and by whole-body MRI (WBMRI) at the Department of Neurology and Neurosurgery of the Hospital Clínico Universidad de Chile (HCUCH), Santiago, Chile. The Ethics Committee of HCUCH and the Chilean National Commission of Scientific Research and Technology (CONICYT) approved the study protocol, and all patients provided informed consent. Twenty-seven patients were included in a previous publication describing the imaging phenotype of dysferlinopathy in WBMRI[13].

Motor disability

Motor disability was determined as described in a previous report[15] using the Motor Function Measure (MFM) scale[18], the Modified Rankin Scale (mRS)[19] and MRC score[20]. The MFM is a 32-item scoring system that provides a numerical measure of the motor capacity of a patient with neuromuscular impairment according to the following 3 functional axes: standing and transfers (D1); axial and proximal motor function (D2); and distal motor function (D3). D1 dimension of the MFM was selected as being the most sensitive to capture the large differences seen in disease severity between patients with dysferlinopathy[15].

MRI and systematic scoring:

Our group previously published the imaging methods used in this study[13]. Fatty infiltration was systematically scored in 61 pairs of muscles according to the scale adapted by Kornblum et al. 2006[16]. Two experienced musculoskeletal radiologists (JD and LS) performed the scoring. They were blinded to patient clinical and genetic data and to the score of the other evaluator. The degree of fatty infiltration was described as follows: none (0), mild (1), moderate (2a), moderate to severe (2b), and severe (3).

Considering the scale as an ordinal parameter for statistical purposes, a moderate infiltration (2a) would be represented as 2, a moderate to severe infiltration as 3, and severe infiltration as 4. Median MRI scores from both evaluators were averaged for individual muscles. Median score for left and right sides was performed to represent the value of those particular muscles. Due to the lack of significant asymmetries in this disorder, averaging between sides results in neither a loss of information nor a bias[12,13].

Heatmaps

To represent the distribution of signal abnormalities in a graphical and understandable manner, heatmaps were built by using R software. Gower's distance was used in the clustering of patients or muscles[16].

Random forests

Random forests were used as a machine-learning approach to study the relationship between distribution of fatty infiltration and clinical features. Random forests (RF) are popular tree-based ensemble machine learning tools that are highly data-adaptive, applicable to "large *p*, small *n*" problems, and able to account for correlation as well as interactions among features[1718]. RF were used to predict three clinical parameters: disease duration, mRS and dimension 1 of MFM. Median fatty infiltration scores for every muscle were used as input parameters. Forests of 1,000 trees were grown using weighted mean-squared error splitting with random input selection to obtain out-of-the-bag estimates for the predicted number of nodes. An algorithm based on minimal depth was used to select a limited number of muscle fatty infiltration scores, that allow a more simple understanding of the relationship[18]. RF were refitted with the obtained variable selection, and the goodness of fit of random forest prediction was calculated by means of out-of-the bag predictions, using Spearman's rho correlation coefficient as indicator. The predictive importance of each variable in each random forests was estimated using

Breiman's permutation variable importance (VIMP). Missing data were imputed using the method published by Ishwaran et al.[19]. Those parameters that were the most important for the prediction were tested with univariate approaches, such as calculating the Spearman's rho coefficient between predicted parameter and infiltration scoring or representing the conditional probability along the possible values of the predicted parameter. Additionally, the adjusted marginal effect of the fatty infiltration score of the most important muscles on the response parameter was plotted (disease duration, mRS, dimension 1 of MFM). RF were implemented using the randomForestSRC R package[20].

Results

Patient characteristics are summarized in Table 1. Imaging features of 33 patients (16 females and 17 males, 9 LGMD2B and 24 MM) from 27 different families were analysed. Figure 1 depicts a heatmap showing individual values of fatty infiltration in every patient. There are three groups of muscles according to fatty infiltration in the cohort. A first group of muscles – **the negative dysferlinopathy fingerprint** – consists of muscles that are relatively spared in most of the patients, even in those severely affected (upper aspect of the heatmap). Head and cervical muscles tend to be spared. At the scapular level, levator scapulae, trapezius and pectoralis minor are also preserved. Other muscles that tend to be uninvolved are popliteus, piriformis, posterior forearm compartment and transversus abdomini. A second group comprises the muscles that are nearly always involved, even in patients with the mildest levels of global infiltration – **the positive dysferlinopathy fingerprint** – (middle aspect of the heatmap). Most of these muscles are from the lower limbs. At the pelvic girdle, gluteus minimus, tensor fasciae latae and obturator externus are commonly involved. At the thigh, quadriceps muscles, semimembranosus, semitendinosus, biceps femori and adductors are usually infiltrated. At the leg, infiltration is common in nearly all the muscles assessed (with the exception of popliteus). Muscles that are part of the positive fingerprint but not from the lower limbs are subscapularis and lumbaris erector spinae. The third group includes muscles in which fatty infiltration varies across the severity spectrum of the disease.

The heatmap clearly shows that fatty infiltration is very heterogeneous in the dysferlinopathy patients of the present study. Random forests were used to explain such heterogeneity. Goodness of fit of the random forests models is sufficiently large for disease duration (Spearman's rho = 0.739), modified Rankin scale (Spearman's rho =

0.723) and dimension 1 of MFM (Spearman's rho = 0.743). Figure 2 shows the selection of variables and their importance in the prediction of each clinical feature.

The most important muscles to predict disease duration are infraspinatus (Spearman's rho between infiltration score and disease duration = 0.777, $p < 0.001$), teres major-minor (rho = 0.781, p-value < 0.001), supraspinatus (rho = 0.774, $p < 0.001$), flexor digitorum longus (rho = 0.634, $p < 0.001$), vastus lateralis (rho = 0.709) and neck extensors (rho = 0.362, $p = 0.038$). Figure 3 shows the conditional probability of an infiltration score equal to or less than 0, 1, 2a, 2b and 3 in these muscles. Figure S1a is a heatmap visualizing the relationship between disease course and every muscle selected in the random forest. Notice how infiltration tends to increase in patients with longer disease duration.

The most important muscles to predict dimension 1 of MFM are teres major-minor (rho= -0.749, $p < 0.001$), triceps (rho = -0.770, $p < 0.001$), sartorius (rho = -0.669, $p < 0.001$), adductor magnus et brevis (rho = -0.669, $p < 0.001$), iliopsoas (rho = -0.718, $p < 0.001$) and gracilis (rho = -0.632, $p < 0.001$). Figure 4 shows the conditional probability of an infiltration score equal to or higher than 0, 1, 2a, 2b and 3 in these muscles. Teres major-minor infiltration progressively increased with increasing disability. Triceps brachii infiltration with score higher or equal to 2a, sartorius infiltration with score higher or equal to 2b or severe infiltration (score 3) of adductor magnus and brevis are infrequent in patients with a MFM-D1 higher than 50%. Figure S1b shows a heatmap of the relationship between MFM-D1 and every muscle selected in the random forest. Note that there are muscles from the three original groups defined in the heatmap involved in disability prediction.

The most important muscles to predict modified Rankin Scale are vastus medialis (rho = 0.752, $p < 0.001$), gracilis (rho = 0.602, $p < 0.001$), infraspinatus (rho = 0.754, $p <$

0.001), sartorius (rho = 0.590, $p < 0.001$), vastus lateralis (rho = 0.735, $p < 0.001$) and vastus intermedius (rho = 0.684, $p < 0.001$). Figure 5 shows the relationship between the infiltration of this muscle and the mRS. Notice that the infiltration of these muscles changes specifically from mRS = 3 (able to walk with assistance) to mRS = 4 (unable to walk without assistance). Figure S1c shows the relationship between mRS and all the muscles that have been selected in the random forests. In general terms, the higher the disability, the more infiltrated all these muscles are. Additionally, the three groups of muscles defined in the global heatmap also emerge when represented according to disability.

Marginal predictions of disease duration (Figure S2), MFM-D1 (Figure S3) and mRS (Figure S4) according to scores of each important muscle reveal a non-linear relationship between such predictions and fatty infiltration. These “breakpoints” are important as they may suggest differential contributions to the prediction of the disease duration at different moments of disease progression.

Discussion.

In a previous report we showed the usefulness of MRI to correlate the degree and distribution of muscle involvement with functional assessment scores[13]. Based on our previous work and studies from other authors ’[12,13], we suggested four characteristic features of muscle involvement in patients with dysferlinopathy. First, there is an early involvement of the posterior compartments of legs and thighs; second there is a marked fatty replacement of gluteus minimus and lumbar erector spinae; third there is an early and selective involvement of the subscapularis, followed in succession by the supraspinatus, deltoid, anterior compartment of the arm and forearm flexors; and fourth, craniofacial muscles are spared. The present work included a larger cohort of patients with dysferlinopathy studied by whole-body MRI, confirming our previous findings and in addition, MRI features were analysed using heatmaps and random forests. This enabled the definition of the range of muscle infiltration with regard to disease severity and disease duration. Representation of muscle infiltration by heatmaps helps to define, in each segment of the body, three types of muscles that may be helpful for MRI diagnosis of dysferlinopathy: muscles that are usually spared (“negative fingerprint”), muscles always impaired (“positive fingerprint”, mostly present in lower limbs); and muscles that are variably affected according to disease status. Taken together, these results provide an objective dataset that leads to the suspicion of disferlinopathy based on the analysis of muscle MRI features. These results confirm, in a separate group of patients, the recently published results of a large international study[21]. Moreover, further analysis by means of a random forest approach enabled the identification of a number of muscles that could be markers of disease duration and disability. The present data show that random forests have a good performance at predicting clinical parameters. The latter means that it is possible to make relatively accurate estimations of the patient’s

functional level or disease duration, based on the degree of fatty infiltration seen in a set of selected muscles. According to the present data, upper limb muscles are more important for the definition of disease duration and disability.

Nevertheless, this method has some limitations. Semi-quantitative subjective scoring is a fast, clinically useful and accessible method to evaluate fatty infiltration, but it is not as accurate as a quantitative approach. A semi-quantitative technique might represent a valid first approach to understand muscle degeneration in dysferlinopathies, but longitudinal studies with quantitative imaging are required to fully validate our results. However, using semi-quantitative scoring is still important for it might drive the design of quantitative studies. For instance, our analysis points out that upper limb muscles (such as infraspinatus and teres major–minor) seem to be more informative than lower limb muscles in order to represent disease duration and disability. The relevance of the infiltration in such muscles has not been evaluated in previous studies[12,13,16] and therefore they might have been overlooked in quantitative approaches.

Infiltration of vastii muscles and medial thigh compartment muscles (gracilis and sartorius) seems to be related to the modified Rankin Scale, suggesting the importance of such muscles in the preservation of walking capacity in dysferlinopathy patients. The latter observation should be further investigated, since it might be particularly important to determine the capacity of the infiltration of these muscles to predict ambulation loss.

As it can be expected in progressive disorders such as muscular dystrophies, disease duration and disability in dysferlinopathy are both related, and there is some overlap in the muscles important for the prediction of these two domains. However, our data show that some muscles are more useful indicators of disease duration, while others are more informative to estimate the time lapse to ambulation loss or impairment of motor abilities. The latter could be explained by the fact that disability and disease duration are

not perfectly correlated (i.e. do not share 100% of their variance). There are some muscles that may be contributing differently to the prediction accuracy of disability and disease duration, because they are related to the portion of the variance that it is not shared by these two parameters.

The use of random forests permits the identification of an objective set of muscles of particular interest for quantitative studies in patients with dysferlinopathies. Simple description of muscle involvement and its representation by using heatmaps may be useful to define the infiltration pattern in myopathies with homogeneous involvement, provided that the sample is large enough and that a wide spectrum of disease severity is explored. However, another approach is needed to identify the muscles that correlate with disease progression. If muscles are variable enough and gradually involved, and the sample size is sufficiently large, supervised learning algorithms may reveal which muscles are the most representative for disease duration and disability. The present work proposes one of the possible data approaches to address this problem, and describes the variability of single muscles pointing to the need for a combined analysis of several muscles to develop a successful biomarker in the future.

Finally, our data seem to indicate that there is a non-linear relationship between disease duration and fatty infiltration. A given degree of fatty infiltration seems to correspond to a particular breakpoint in the course of the disease. This may be relevant when trying to interpret muscle MRI, and may be used as a reference value to describe muscle infiltration during the disease course. Breakpoints may be used to simplify interpretation, e.g. a fibroadipose score equal to or higher than 2a in infraspinatus is indicative of advanced disease. For example, if there is suspicion of dysferlinopathy in a given patient due to clinical features or genetic results, and we decide to compare the patient's muscle MRI with the known pattern of dysferlinopathy, our interpretation of

the infraspinatus infiltration should take into account the patient's point in the disease trajectory, as the infraspinatus would be infiltrated if the patient has a long disease duration, but it might be preserved in patients with short disease duration. There are several possible explanations for such non-linear relationships: a non-homogeneous pace of progression of fatty infiltration in the natural history of the disease; a variable distribution of atrophy in different parts of the body due to mobility loss or contractures during the disease course, or simply as a consequence of using semi-quantitative ordinal scoring systems.

In conclusion, the use of heatmaps and machine-learning techniques enables a better definition of the imaging pattern of dysferlinopathy and identification of key muscles to follow the progression of disease objectively and reliably. We believe that the data presented here will contribute to the diagnosis of dysferlinopathy, as well as to the follow-up of the natural course of the disease. Moreover, by means of integrating this information into the design of quantitative MRI studies, eventual modifications of the disease natural course by future therapeutic interventions may be sucessfully quantified.

List of abbreviations

DYSF Dysferlin

LGMD2B Limb girdle muscular dystrophy 2B

MM Miyoshi myopathy

mRS Modified Rankin's scale

MFM Motor Function Measure

MFM-D1 Motor Function Measure Dimension 1

MRI Magnetic resonance imaging

RF Random forests

WBMRI Whole-body magnetic resonance imaging

References

1. Bashir R, Britton S, Strachan T, Keers S, Vafiadaki E, Lako M, Richard I, Marchand S, Bourg N, Argov Z, Sadeh M, Mahjneh I, Marconi G, Passos-Bueno MR, Moreira Ede S, Zatz M, Beckmann JS, Bushby K. A gene related to Caenorhabditis elegans spermatogenesis factor fer-1 is mutated in limb-girdle muscular dystrophy type 2B. Nature genetics 1998;20(1):37-42.
2. Liu J, Aoki M, Illa I, Wu C, Fardeau M, Angelini C, Serrano C, Urtizberea JA, Hentati F, Hamida MB, Bohlega S, Culper EJ, Amato AA, Bossie K, Oeltjen J, Bejaoui K, McKenna-Yasek D, Hosler BA, Schurr E, Arahata K, de Jong PJ, Brown RH, Jr. Dysferlin, a novel skeletal muscle gene, is mutated in Miyoshi myopathy and limb girdle muscular dystrophy. Nature genetics 1998;20(1):31-36.
3. Aoki M, Liu J, Richard I, Bashir R, Britton S, Keers SM, Oeltjen J, Brown HE, Marchand S, Bourg N, Beley C, McKenna-Yasek D, Arahata K, Bohlega S, Cupler E, Illa I, Majneh I, Barohn RJ, Urtizberea JA, Fardeau M, Amato A, Angelini C, Bushby K, Beckmann JS, Brown RH, Jr. Genomic organization of the dysferlin gene and novel mutations in Miyoshi myopathy. Neurology 2001;57(2):271-278.
4. Illa I, Serrano-Munuera C, Gallardo E, Lasa A, Rojas-Garcia R, Palmer J, Gallano P, Baiget M, Matsuda C, Brown RH. Distal anterior compartment myopathy: a dysferlin mutation causing a new muscular dystrophy phenotype. Annals of neurology 2001;49(1):130-134.
5. Barthelemy F, Wein N, Krahn M, Levy N, Bartoli M. Translational research and therapeutic perspectives in dysferlinopathies. Molecular medicine 2011;17(9-10):875-882.
6. Park HJ, Hong JM, Suh GI, Shin HY, Kim SM, Sunwoo IN, Suh BC, Choi YC. Heterogeneous characteristics of Korean patients with dysferlinopathy. Journal of Korean medical science 2012;27(4):423-429.

7. Cupler EJ, Bohlega S, Hessler R, McLean D, Stigsby B, Ahmad J. Miyoshi myopathy in Saudi Arabia: clinical, electrophysiological, histopathological and radiological features. Neuromuscular disorders : NMD 1998;8(5):321-326.
8. Katz JS, Rando TA, Barohn RJ, Saperstein DS, Jackson CE, Wicklund M, Amato AA. Late-onset distal muscular dystrophy affecting the posterior calves. Muscle & nerve 2003;28(4):443-448.
9. Ro LS, Lee-Chen GJ, Lin TC, Wu YR, Chen CM, Lin CY, Chen ST. Phenotypic features and genetic findings in 2 chinese families with Miyoshi distal myopathy. Archives of neurology 2004;61(10):1594-1599.
10. Brummer D, Walter MC, Palmbach M, Knirsch U, Karitzky J, Tomczak R, Braun C, Grundhoff N, Bornemann A, Muller CR, Lochmuller H, Schreiber H. Long-term MRI and clinical follow-up of symptomatic and presymptomatic carriers of dysferlin gene mutations. Acta myologica : myopathies and cardiomyopathies : official journal of the Mediterranean Society of Myology 2005;24(1):6-16.
11. Illa I, De Luna N, Dominguez-Perles R, Rojas-Garcia R, Paradas C, Palmer J, Marquez C, Gallano P, Gallardo E. Symptomatic dysferlin gene mutation carriers: characterization of two cases. Neurology 2007;68(16):1284-1289.
12. Paradas C, Llauger J, Diaz-Manera J, Rojas-Garcia R, De Luna N, Iturriaga C, Marquez C, Uson M, Hankiewicz K, Gallardo E, Illa I. Redefining dysferlinopathy phenotypes based on clinical findings and muscle imaging studies. Neurology 2010;75(4):316-323.
13. Diaz J, Woudt L, Suazo L, Garrido C, Caviedes P, AM CA, Castiglioni C, Bevilacqua JA. Broadening the imaging phenotype of dysferlinopathy at different disease stages. Muscle & nerve 2016;54(2):203-210.
14. Jin S, Du J, Wang Z, Zhang W, Lv H, Meng L, Xiao J, Yuan Y. Heterogeneous characteristics of MRI changes of thigh muscles in patients with dysferlinopathy. Muscle & nerve 2016;54(6):1072-1079.
15. Woudt L, Di Capua GA, Krahn M, Castiglioni C, Hughes R, Campero M, Trangulao A, Gonzalez-Hormazabal P, Godoy-Herrera R, Levy N, Urtizberea JA, Jara L, Bevilacqua JA. Toward an objective measure of functional disability in dysferlinopathy. Muscle & nerve 2016;53(1):49-57.
16. Gower JC. A General Coefficient of Similarity and Some of Its Properties. Biometrics 1971;27(4):857-871.
17. Chen X, Ishwaran H. Random forests for genomic data analysis. Genomics 2012;99(6):323-329.
18. Ishwaran H, Kogalur UB, Gorodeski EZ, Minn AJ, Lauer MS. High-Dimensional Variable Selection for Survival Data. Journal of the American Statistical Association 2010;105(489):205-217.
19. Ishwaran H, Kogalur UB, Blackstone EH, Lauer MS. Random survival forests. The annals of applied statistics 2008:841-860.
20. Ishwaran H, Kogalur UB, Kogalur MUB. Package 'randomForestSRC'2017.
21. Diaz-Manera J, Fernandez-Torron R, J LL, James MK, Mayhew A, Smith FE, Moore UR, Blamire AM, Carlier PG, Rufibach L, Mittal P, Eagle M, Jacobs M, Hodgson T, Wallace D, Ward L, Smith M, Stramare R, Rampado A, Sato N, Tamaru T, Harwick B, Rico Gala S, Turk S, Coppenrath EM, Foster G, Bendahan D, Le Fur Y, Fricke ST, Otero H, Foster SL, Peduto A, Sawyer AM, Hilsden H, Lochmuller H, Grieben U, Spuler S, Tesi Rocha C, Day JW, Jones KJ, Bharucha-Goebel DX, Salort-Campana E, Harms M, Pestronk A, Krause S, Schreiber-Katz O, Walter MC, Paradas C, Hogrel JY, Stojkovic T, Takeda S, Mori-Yoshimura M, Bravver E, Sparks S, Bello L, Semplicini

C, Pegoraro E, Mendell JR, Bushby K, Straub V, Jain COSC. Muscle MRI in patients with dysferlinopathy: pattern recognition and implications for clinical trials. Journal of neurology, neurosurgery, and psychiatry 2018;89(10):1071-1081.

TABLES

Table 1. Characteristics of the 33 patients with dysferlinopathy.

Patient ID	Family	Sex	Age (yrs)	Phenotype at onset	Disease duration (yrs)	Ambulatory status	MFM-D1	mRS
P8	7*	F	31	LGMD2B	5	A	79.5	2
P11	11*	F	23	LGMD2B	13	A	43.6	4
P14	14	F	43	LGMD2B	13	W	18.0	4
P17	17	F	26	LGMD2B	10	C	25.6	4
P18	18	M	19	LGMD2B	4	A	38.5	3
P27	27	F	32	LGMD2B	15	C	23.1	4
P36	33*	F	38	LGMD2B	8	A	48.7	3
C03	28	F	37	LGMD2B	15	W	15.4	4
C05	31	M	45	LGMD2B	12	C	10.3	4
P1	1	M	34	MM	14	W	7.7	5
P4	4	M	22	MM	7	A	59.0	2
P5	5	F	30	MM	12	C	48.7	4
P7	7*	F	28	MM	8	C	25.6	4
P9	9	M	24	MM	12	W	18.0	5
P12	11*	M	26	MM	6	A	61.5	3
P13	13	M	43	MM	18	C	18.0	4
P15	15*	M	24	MM	7	A	56.4	3
C06	15*	M	25	MM	7	A	58.9	3
P16	16	M	28	MM	8	A	74.4	3
P19	19*	F	37	MM	24	C	NA	4
P20	19*	M	39	MM	20	C	NA	4
P22	19*	M	34	MM	19	W	0	5
P23	23	F	20	MM	3	A	82.1	2
P24	24	M	29	MM	10	A	38.5	4
P25	25	F	51	MM	18	C	35.9	4
P32	32	M	32	MM	5	A	35.9	3
P33	33*	M	30	MM	10	A	30.8	3
P34	34	F	27	MM	5	A	74.4	4
P39	39	F	29	MM	5	A	69.3	3
P44	44	F	30	MM	7	A	53.9	3
C01	29	M	20	MM	7	A	71.8	3
C02	30	M	31	MM	7	W	43.6	4
C04	35	F	33	MM	19	C	38.5	3

Family identification and patient identification starting with "P" correspond to patients reported in Would et al. (15). Patients with other member in the cohort are indicated with an asterisk in the family ID.
M: Male. F: Female. MM: Miyoshi Myopathy. LGMD2B: Limb girdle muscular dystrophy 2B. A: Ambulant, C: Walk with cane and W: Wheelchair bounded. MFM-D1: Motor Function Measure Dimension D1. mRS: Modified Rankin Scale.

FIGURE LEGENDS

Figure 1. Heatmap representing scores for T1 signal in dysferlinopathy patients. A dendrogram above the heatmap represents the similarities between patients (the closer the junction to the bottom, the more similarity between two given patients). Muscles are also ordered according to hierarchical clustering with the dendrogram shown at the right aspect of the figure. Shading represents the intensity of muscle changes (see legend in the left upper aspect of the figure).

Figure 2. Variable importance in the random forest models. MFM: Motor Function Measure.

Figure 3. Cumulative probability of having a value of fatty infiltration below some cut-off value as a function of disease duration.

Figure 4. Cumulative probability of having a value of fatty infiltration higher than some cut-off value as a function of MFM dimension 1. MFM: Motor Function Measure.

Figure 5. Cumulative probability of having a value of fatty infiltration below some cutting values depending on modified Rankin scale.