

Predictors of functional outcomes in patients with facioscapulohumeral muscular dystrophy

 Natalie K. Katz,¹ John Hogan,² Ryan Delbango,² Colin Cernik,³ Rabi Tawil⁴ and Jeffrey M. Statland⁵

See Alfano and Mozaffar (doi:10.1093/brain/awab389) for a scientific commentary on this article.

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most prevalent muscular dystrophies characterized by considerable variability in severity, rates of progression and functional outcomes. Few studies follow FSHD cohorts long enough to understand predictors of disease progression and functional outcomes, creating gaps in our understanding, which impacts clinical care and the design of clinical trials. Efforts to identify molecularly targeted therapies create a need to better understand disease characteristics with predictive value to help refine clinical trial strategies and understand trial outcomes.

Here we analysed a prospective cohort from a large, longitudinally followed registry of patients with FSHD in the USA to determine predictors of outcomes such as need for wheelchair use. This study analysed de-identified data from 578 individuals with confirmed FSHD type 1 enrolled in the United States National Registry for FSHD Patients and Family members. Data were collected from January 2002 to September 2019 and included an average of 9 years (range 0–18) of follow-up surveys. Data were analysed using descriptive epidemiological techniques, and risk of wheelchair use was determined using Cox proportional hazards models. Supervised machine learning analysis was completed using Random Forest modelling and included all 189 unique features collected from registry questionnaires. A separate medications-only model was created that included 359 unique medications reported by participants.

Here we show that smaller allele sizes were predictive of earlier age at onset, diagnosis and likelihood of wheelchair use. Additionally, we show that females were more likely overall to progress to wheelchair use and at a faster rate as compared to males, independent of genetics. Use of machine learning models that included all reported clinical features showed that the effect of allele size on progression to wheelchair use is small compared to disease duration, which may be important to consider in trial design. Medical comorbidities and medication use add to the risk for need for wheelchair dependence, raising the possibility for better medical management impacting outcomes in FSHD.

The findings in this study will require further validation in additional, larger datasets but could have implications for clinical care, and inclusion criteria for future clinical trials in FSHD.

1 Department of Neurology, Children's Mercy Hospital, Kansas City, MO 64108, USA

2 Department of Artificial Intelligence, Albytes, LLC, Hurley, NY 12443, USA

3 Department of Biostatistics, University of Kansas Medical Center, Kansas City, KS 66160, USA

4 Department of Neurology, University of Rochester Medical Center, Rochester, NY 14642, USA

5 Department of Neurology, University of Kansas Medical Center, Kansas City, KS 66160, USA

Correspondence to: Jeffrey M. Statland, MD
 Department of Neurology, The University of Kansas Medical Center
 3901 Rainbow Blvd, MS 2012, Kansas City, KS 66160, USA
 E-mail: JStatland@kumc.edu

Keywords: facioscapulohumeral muscular dystrophy; machine learning; artificial intelligence; wheelchair use; functional outcomes

Abbreviations: FSHD = facioscapulohumeral muscular dystrophy; SHAP = Shapley Additive Explanation; The Registry = US National Registry of Myotonic Dystrophy and FSHD Patients and Family Members

Introduction

Facioscapulohumeral muscular dystrophy (FSHD) is one of the most prevalent types of muscular dystrophy^{1–4} with considerable variability in age at onset, rates of progression and motor outcomes.^{5–8} Cross-sectional and limited longitudinal studies have shown an association between the D4Z4 repeat size, age at diagnosis, age of wheelchair use, and overall severity of disease.^{9–12} Outcome measurements in most studies have largely focused on muscle strength testing, clinical severity score ratings,¹³ and more recently MRI of muscle mass,^{14,15} but no study has linked these measurements to motor outcomes such as wheelchair use or need for non-invasive ventilation. Currently, there are few longitudinal studies available with follow-up data of sufficient duration to inform our understanding of what drives this clinical variability and motor outcomes, creating a gap in our understanding of FSHD, which impacts care decisions and clinical trial design.

We previously analysed data from a smaller cohort of subjects from the United States National Registry of FSHD Patients and Family Members (hereafter referred to as 'The Registry'), which contains over 180 data elements.¹² However, a majority of this information could not be analysed for contributions to functional outcomes due to the volume and complexity of the data. Machine learning is a powerful technology that offers the potential to better understand predictors of outcomes in FSHD in large datasets by allowing analysis of large and complex datasets. A similar analysis based on machine learning of a large dataset of patients with Huntington's disease was used by Sun *et al.*¹⁶ to model disease progression. The investigators were able to identify numerous disease states and predict the rate of transition from one state to the next, something that has been historically difficult to elucidate due to the slowly progressive nature of the disease. We posited that use of machine learning with The Registry might help better understand disease progression over time by identifying factors that contribute to the clinical heterogeneity and functional outcomes seen in patients with FSHD.

Our study used a combination of traditional epidemiological methods and supervised machine learning techniques to analyse an average of 9 years of longitudinally collected data from participants enrolled in The Registry. The primary outcome in this study was progression to wheelchair use. Machine learning technology allowed us to incorporate all of the data elements collected in The Registry, including information on medical comorbidities and medication use.

Materials and methods

Study design

We analysed a prospective cohort study using de-identified data collected on individuals participating in The Registry located at the

University of Rochester Medical Center in Rochester, New York (<https://www.urmc.rochester.edu/neurology/national-registry.aspx>). Enrolment in The Registry is voluntary, and all individuals give informed consent for de-identified data to be used in future research projects on FSHD at the time of enrolment. Data analysed were collected from January 2002 to September 2019. This study was determined to be 'not a human subject' research after IRB review because the data were already collected and de-identified. The study protocol was submitted to an independent Registry Scientific Advisory Board who approved the proposed study.

Patient population

Participants were included for analysis if they were confirmed to have FSHD type 1 (FSHD1) based on positive genetic testing, which was defined by The Registry as confirmation of: (i) commercial testing based on medical chart review at entry into The Registry; (ii) commercial testing in a first degree relative based on chart review; or (iii) research testing performed at Leiden University Medical Center as part of a FSHD research study in conjunction with The Registry. Participants were excluded if they were not genetically confirmed to have FSHD1. The number of patients with FSHD type 2 (FSHD2) was too low to allow sufficient power for the analysis, therefore, these patients were excluded.

Outcomes

At the time of enrolment in The Registry, a limited chart review is conducted by staff at The Registry to verify the presence of clinical features consistent with FSHD (facial, shoulder and/or foot dorsiflexion weakness) on prior exams; to review genetic testing for molecular classification of disease type (FSHD1 or FSHD2 or other); and to classify the likelihood of a diagnosis of FSHD (clinically definite, possible, unaffected with blood relative, or not FSHD). Once enrolled, individuals fill out annual surveys that include detailed questions related to current functional abilities, signs and symptoms (such as patient reported age at diagnosis and symptom onset), medical comorbidities and medication use. Data analysed in this study contained 189 unique features (independent variables) and included information such as (but not limited to): gender, year of birth, diagnostic tests, genetic analysis, presenting symptoms, family history, employment, education, use of assistive devices, medical comorbidities, and medication use. Examples of these questionnaires can be found at: <https://www.urmc.rochester.edu/neurology/national-registry/join.aspx>. This prospective information was used to identify factors that might contribute to disease progression using a combination of traditional epidemiological methods and supervised machine learning techniques.

The primary outcome for the machine learning analysis was progression to wheelchair use for distance and/or full-time use,

minimizing the potential for recall bias. The Registry contains several types of data including nominal, count, and continuous variables that were included for machine learning analysis. Nominal variables primarily consisted of ‘yes or no’ answers to questions related to medical comorbidities (such as breathing concerns, arthritis, pneumonia, hypertension, cardiac concerns, and constipation), presenting symptoms (such as facial weakness or proximal upper extremity weakness), and types of medications used (such as vitamins, minerals or non-steroidal anti-inflammatory drugs). Gender is included in this as well. Count and continuous variables included number of medications, age at diagnosis, current age of the patient, and initial age at symptom onset. Some categories required calculation prior to inclusion in the analysis, such as disease duration (defined as the patient’s current age minus the age of initial symptom onset), and the length of time the patient spent undiagnosed (defined as age at diagnosis minus the age of initial symptoms onset). Body mass index (BMI) was determined using standard calculations. All 189 features collected from The Registry questionnaires were included for development of the machine learning model and used to assess features that might be predictive of progression to wheelchair use.

Registry data captured the D4Z4 small allele size in kilobytes from commercial testing. For analysis, this small allele size was converted into estimated number of D4Z4 repeats to be consistent with the literature using the following formula: (small allele size – 6) / 3.3.

Statistical analysis

Baseline population characteristics were analysed using descriptive techniques including means and standard deviation (SD), median and interquartile ranges (IQR), or counts and frequencies according to the data type. Rudimentary categorical analysis methods such as Fisher’s exact test were used to compare outcomes across different demographic groups. Log Rank analysis was used to assess age of wheelchair use based on gender with one degree of freedom. Log Rank analysis was also used to assess median time from age at diagnosis to age of wheelchair use based on gender (one degree of freedom), genetics (D4Z4 repeat length; two degrees of freedom) and gender + genetics (seven degrees of freedom). The dependent variable was progression to wheelchair use, and gender and genetics were the independent variables. One-way Cox proportional hazards models were used to examine the hazard ratios over time for gender and D4Z4 repeat length (independent variables), using age as the timescale. Progression to wheelchair use was the dependent variable. The hazard ratio was determined by taking the log odds of the coefficients, and 95% confidence intervals (CI) were taken from the final model. The test for proportional hazards was used to assess the assumption of proportional hazards when using the models. For participants who were not genetically defined but had a first-degree relative who was genetically defined, the allele size of the relative was substituted for the participant in the analysis. Early genetic testing was reported as only ‘positive’ or ‘negative’ and therefore some individuals do not have a kilobyte allele size reported. Where appropriate, individuals who were not genetically defined and/or did not have a kilobyte size reported were excluded from model analysis. An alpha level less than $P = 0.05$ was considered statistically significant. All statistical analyses were performed using R programming and statistical software (version v4.0.4).

Python and Scikit-learn were utilized to develop the supervised machine learning model. Random Forest was selected for model development given its success in similar projects for disease modelling.^{16–20} Characteristics used to select the best model included accuracy and area under the curve (AUC). For model

accuracy, an AUC approaching 0.8 is considered good, and 0.9 considered very good.²¹ To validate the best fit model, 15 records were randomly extracted from the annual updates: 12 ‘test subjects’ who progressed from walking to wheelchair use, and three ‘controls’ who did not progress to wheelchair use. The model correctly predicted the outcomes of these individuals. Once validated, the model was run against the entire dataset to identify features that were predictive of wheelchair use. Features with a relative importance value >0.03 were felt to have a significant contribution towards influencing wheelchair use.

Shapley Additive Explanation (SHAP) was also used in an effort to explain the relative importance of each feature both locally and globally to wheelchair use, as predicted by the machine learning model.²² SHAP assigns points on a scale of –1.0 to 1.0, with a value of 1.0 representative of a feature’s influence towards wheelchair use and a value of –1.0 representative of a feature’s influence towards no wheelchair use.²² Each dot represents one individual in the dataset. Red dots represent an individual having a value within that feature that is proportionally higher than the median value; blue dots represent an individual having a value that is proportionally lower than the median value; and purple dots represent an individual having a value that is proportionally similar to the median value. Binary categorical features within the set including gender, medical comorbidities and medication use only had two possible values, 0 and 1, so no purple dots could be represented by those attributes. Attributes with a wider x-axis have a stronger influence on prediction of wheelchair use.

Medication-only model

A separate model containing only medications was created to evaluate the influence of specific categories of medications on wheelchair use; all other data were stripped from the model. Names of medications were normalized and categorized based on medication class.

Data availability

All de-identified data used in this study are available upon request from The Registry (<https://www.urmc.rochester.edu/neurology/national-registry.aspx>). Python software was used to generate the Random Forest model and is openly available at <https://scikit-learn.org/stable/>.

Results

At the time data were received from The Registry, a total of 1030 participants were enrolled. Seven patients with FSHD2 and 445 participants who were not genetically defined were excluded, leaving a total of 578 participants with FSHD1 who were genetically defined and were included for analysis. Data analysed were collected from January 2002 to September 2019. An average of 9 years of follow-up data were analysed, with a range from 0 to 18 years. The majority of registry participants were Caucasian, with a slight male predominance (Table 1). Most participants were middle-aged, but ages spanned the lifespan (range 11–100). Over half of participants had achieved a college degree or higher (63.9%), and most were employed at enrolment (52.6%). On average, there was a 13-year delay between symptom onset and diagnosis (Table 1). Over half of participants ($n = 320$, 55%) reported symptom onset prior to age 18 (data not shown). Most participants (58.8%) had 4–7 D4Z4 repeats; 20.6% of participants had 8–10 D4Z4 repeats, and 10.4% of participants had 1–3 D4Z4 repeats. The five most common initial symptoms reported are shown (Table 1). Females were more likely than males to report facial weakness as their initial

Table 1 Registry demographics and disease characteristics

	Female	Male	Overall
Demographics			
Gender (% overall)	277 (47.9%)	301 (52.1%)	578 (100%)
Current age, median (1st, 3rd quartile) ^a	59 (45.5, 70.5)	59 (46, 69)	59 (46, 70)
Race (% overall)			
Caucasian	250 (90.3%)	275 (91.4%)	535 (92.6%)
Black	0 (0%)	4 (1.3%)	4 (0.7%)
Asian	8 (2.9%)	7 (2.3%)	15 (2.6%)
Other/not reported	6 (2.2%)	4 (1.3%)	20 (3.5%)
Ethnicity (% overall)			
Hispanic/Latino	11 (4%)	9 (3%)	20 (3.5%)
Education (% overall)			
College, Masters or Doctoral degree	170 (61.4%)	199 (66.1%)	369 (63.9%)
Technical school	17 (6.1%)	20 (6.6%)	37 (6.4%)
Elementary + High school	84 (30.3%)	74 (24.6%)	158 (27.3%)
Missing	6 (2.2%)	8 (2.7%)	14 (2.4%)
Employed (% overall) ^b			
Yes	129 (47.1%)	172 (57.7%)	301 (52.4%)
Disease characteristics			
Age at diagnosis, median (1st, 3rd quartile) ^c	31.5 (16, 46)	29 (18, 47)	30.0 (18, 47)
Age of initial symptoms, median (1st, 3rd quartile) ^d	16 (10, 31)	18 (13, 27)	17 (12, 29.5)
D4Z4 repeat length ^e			
1–3	37	23	60 (10.4%)
4–7	175	165	340 (58.8%)
8–10	43	76	119 (20.6%)
Initial symptoms			
Facial weakness	72	24	96 (17%)
Upper extremity, proximal weakness	77	105	182 (31%)
Upper extremity, unspecified weakness	8	19	27 (5%)
Lower extremity, unspecified weakness	27	29	56 (9.7%)
Atrophy, muscle mass change	5	25	30 (5%)
Use of a wheelchair (% overall)			
Yes	80	57	137 (23.7%)
Breathing problems (% overall)			
Yes	58 (20.9%)	47 (15.6%)	105 (18.2%)
Breathing problems due to FSHD (% overall)			
Yes	23 (8.3%)	21 (7%)	44 (7.6%)
Use a breathing machine (% overall)			
Yes	4 (1.4%)	24 (8%)	28 (4.8%)

^aTwo participants (one male, one female) did not report birth year. Current age calculated based on birth year as of 31 December 2020.

^bSix participants were excluded due to age < 16 years as they were not considered old enough to be a part of the workforce.

^c12 missing; four females, eight males.

^d31 missing; 16 females, 15 males.

^e59 missing; 22 female, 37 male.

presenting symptom, whereas males were more likely to report upper extremity weakness and muscle atrophy as their initial presenting symptoms. Most participants (76.3%) were ambulatory at the time of enrolment. Regarding breathing difficulty, 7.6% of participants reported breathing difficulty due to FSHD and 4.8% reported use of a breathing machine.

When we looked at repeat length and gender, we found a higher frequency of females in the 1–3 D4Z4 repeat category ($P = 0.002$), males in the 8–10 D4Z4 repeat group ($P = 0.002$), and a roughly equal distribution of males and females in the 4–7 D4Z4 repeat category (Table 1). Consistent with previous studies, there was a relationship between D4Z4 repeat length and the age at diagnosis, with individuals with 1–3 repeats diagnosed at a markedly younger age compared to those with higher numbers of repeats (Fig. 1A). Overall, there was not a significant difference in age at diagnosis between males and females with 1–3 or 8–10 D4Z4 repeats; however, females with 4–7 repeats were diagnosed ~10 years later than males in this category (Fig. 1B). When

comparing genetics, age at diagnosis, and presenting symptoms, individuals with 1–3 repeats (10–18 kb allele size) were most likely to report facial weakness (53.7%) as their initial symptom while all others were most likely to report proximal upper extremity weakness as their initial symptom (Fig. 2).

At the time of enrolment in The Registry, 137 individuals reported using a wheelchair. Kaplan-Meier estimates of these individuals showed a median age of wheelchair use for individuals with 1–3 repeats of 14 years (95% CI: 13, 38); 46 years for individuals with 4–7 repeats (95% CI: 44, 52); and 60 years for individuals with 8–10 repeats (95% CI: 55, 68; data not shown). These results are consistent with our previously reported study¹² using a similar cohort, and also in line with other reported ages in the literature.⁹ Of the 441 individuals at risk for wheelchair use, 286 progressed to using a wheelchair giving an incidence estimate of 0.65. Across all groups, we found that females were significantly more likely than males to use a wheelchair ($P = 0.003$; Fig. 3A and Table 2), and that females have a shorter time from age at diagnosis to wheelchair

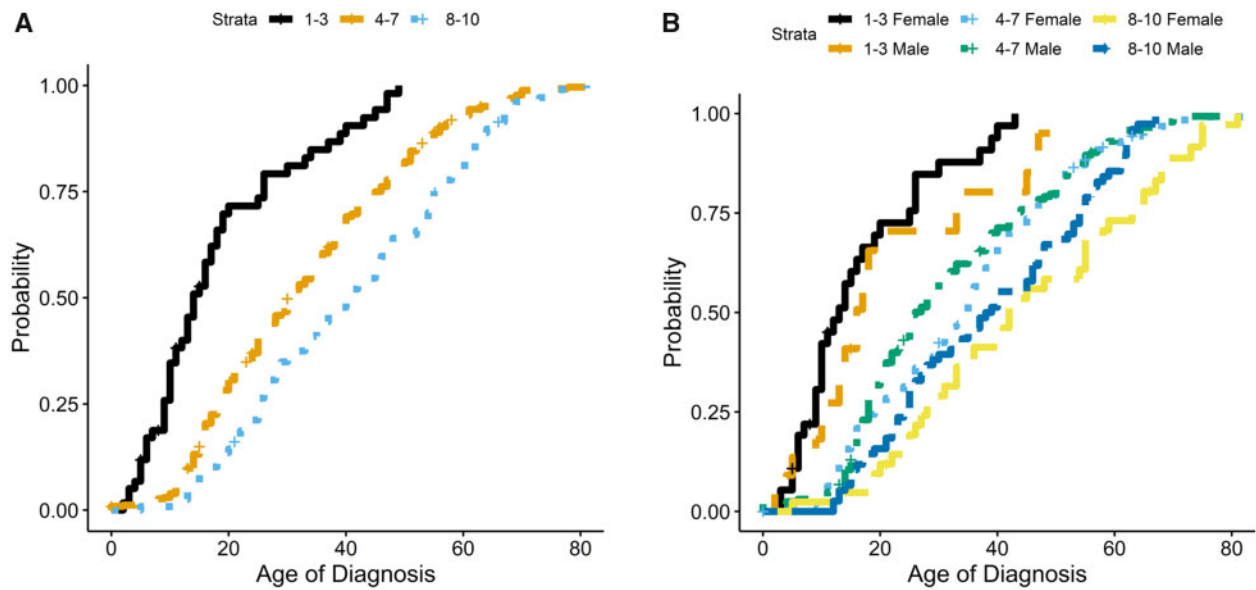


Figure 1 Repeat length, age at diagnosis and gender. Cumulative probability plots were used to compare repeat length and age at diagnosis (A), as well as repeat length and age at diagnosis with respect to gender (B). (A) A median age of diagnosis of 14 (95% CI: 11, 17) for all individuals with 1–3 repeats; 30 years (95% CI: 27, 34) for all individuals with 4–7 repeats; and 40 years (95% CI: 35, 46) for all individuals with 8–10 repeats. When separated by gender (B), there does appear to be a separation in the age at which males and females with 4–7 repeats were diagnosed. Males had a median age of diagnosis of 25 (95% CI: 24, 30) whereas females had a median age of diagnosis of 35 (95% CI: 30, 37). There is no difference in the median age at diagnosis for males and females with 1–3 or 8–10 repeats. Females in the 1–3 repeat category were diagnosed at a median age of 11 (95% CI: 10, 17) whereas males were diagnosed at a median age of 16 (95% CI: 13, 33). Females in the 8–10 repeat category were diagnosed at a median age of 42 (95% CI: 33, 55) whereas males were diagnosed at a median age of 38.5 (95% CI: 32, 47).

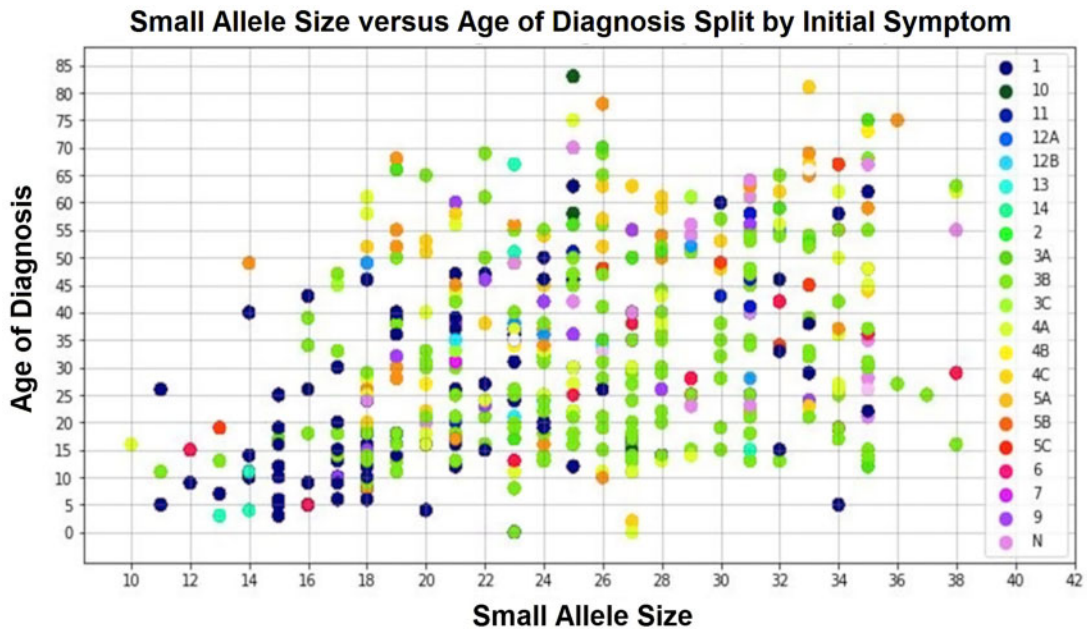


Figure 2 Small allele size (repeat length), age at diagnosis and initial symptoms. When comparing initial presenting symptoms to age at diagnosis and repeat length, we see a cluster of facial weakness (1 = dark blue dots) in individuals with the smallest repeat lengths (1–3 repeats = 10–18 kb allele size). We also see a cluster of proximal upper extremity weakness (3A = green dots) in individuals with medium (4–7 repeats; 18–30 kb allele size) repeat lengths. *Right axis:* 1 = facial weakness; 2 = trunk weakness; 3A = proximal upper extremity (UE) weakness; 3B = distal UE weakness; 3C = unspecified UE weakness; 4A = proximal lower extremity (LE) weakness; 4B = distal LE weakness; 4C = unspecified LE weakness; 5A = pain in the back/trunk; 5B = pain in the UE; 5C = pain in the LE; 6 = fatigue or generalized weakness; 7 = atrophy or loss of muscle mass; 8 = muscle cramps; 9 = abnormal laboratory values; 10 = family history of FSHD; 11 = no symptoms; 12A = injury due to falling; 12B = injury not due to falling; 13 = unable to classify; 14 = sensory changes; N = missing data.

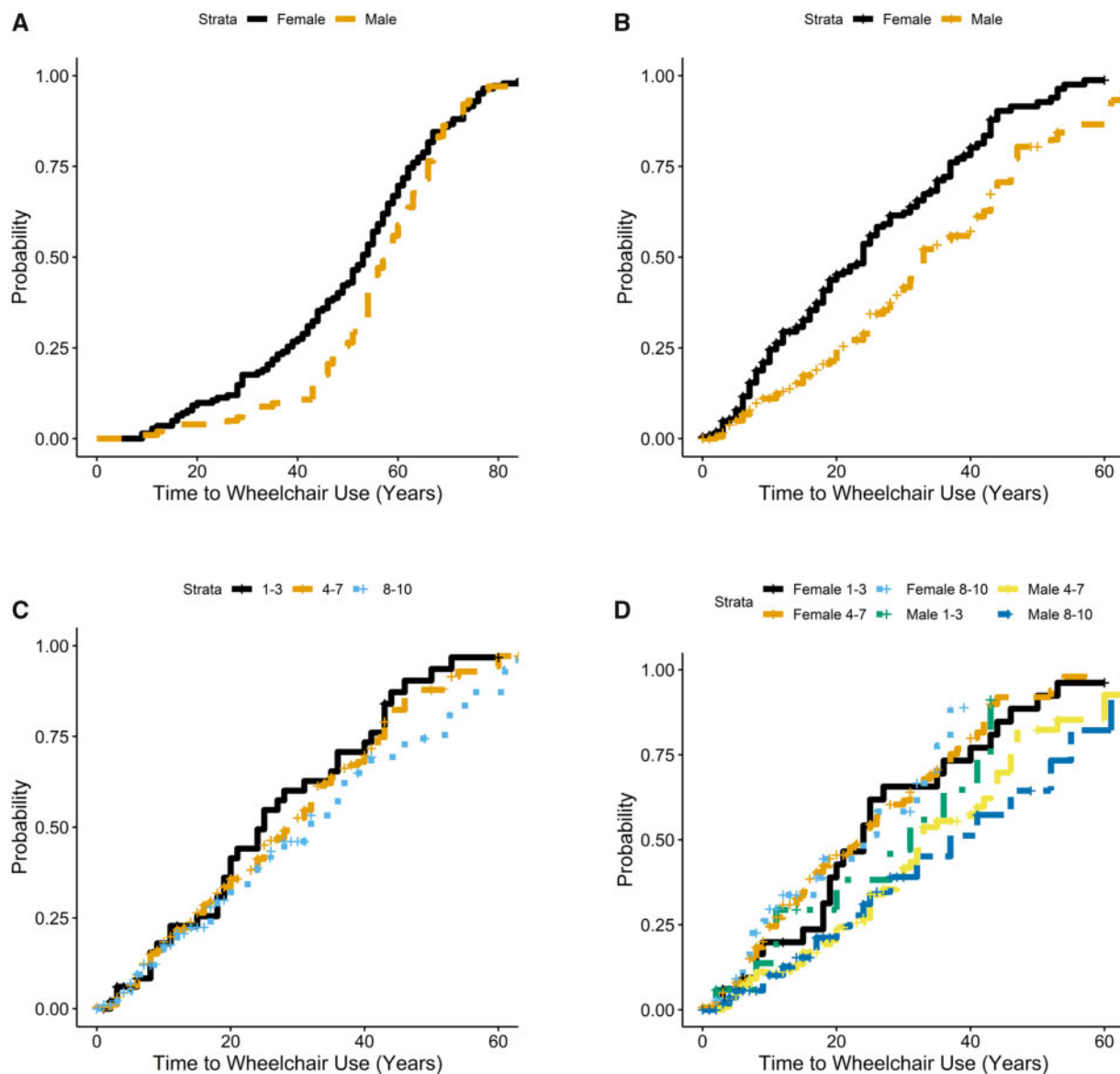


Figure 3 Females are more likely to progress to wheelchair use overall. Log-rank analysis of baseline and longitudinal data shows that (A) females were more likely at all ages to use a wheelchair compared to males, with a median age of wheelchair use of 59 (95% CI: 56, 62) whereas males had a median age of wheelchair use at 64 (95% CI: 62, 68). Females have a shorter length of time from age at diagnosis to age of wheelchair use (B), with a median difference of 23 years for females (95% CI: 19, 26) and 32 years for males (95% CI: 29, 37). There is no significant difference between D4Z4 repeat length and length of time from diagnosis to wheelchair use ($P = 0.2$) (C), but when separated by gender we again see that females were more likely than males to progress to wheelchair use ($P = 4 \times 10^{-4}$) (D). Females in the 1–3 D4Z4 repeat category have a median time of progression to wheelchair use of 23 years (95% CI: 15, 31) whereas males with 1–3 D4Z4 repeats have a median time of progression to wheelchair use of 28 years (95% CI: 20, n/a). Females in the 4–7 D4Z4 repeat category had a median time of progression to wheelchair use of 22 years (95% CI: 18, 27) whereas males had a median time of progression of 33 years (95% CI: 29, 40). Females in the 8–10 D4Z4 repeat category had a median time of progression to wheelchair use of 20 years (95% CI: 12, n/a) whereas males had a median time of progression of 28 years (95% CI: 18, 54).

use ($P = 1 \times 10^{-6}$; Fig. 3B), even after adjusting for differences in allele length (Fig. 3C and D). Using all longitudinal data, we see that individuals with 1–3 D4Z4 repeats were significantly more likely than those with 4–7 repeats to use a wheelchair (Table 2). Individuals with 8–10 repeats were less likely than those with 4–7 repeats to progress to wheelchair use.

Machine learning analysis

The Random Forest model selected had an accuracy of 0.79 and AUC of 0.85 for predicting wheelchair use. Both the Random Forest model and SHAP analysis found that age-related features had a

high predictive value as to whether or not someone will progress to wheelchair use, with longer disease duration, older current age of the patient, younger age at diagnosis and younger age at symptom onset (Fig. 4A and B). The feature found to have the second highest influence on progression to wheelchair use was the number of medications participants reported taking, with higher numbers of medication associated with wheelchair use. The presence of several medical comorbidities were predicted to increase risk for progression to wheelchair use, including breathing problems, pneumonia, arthritis, constipation, heart problems and psychiatric concerns. Female gender was associated with higher likelihood of wheelchair use. Individuals with a BMI lower than the median at

Table 2 Wheelchair use

	Overall (n = 578)	Wheelchair use	HR (95% CI)
Gender			
Female	277	80	1.44 (1.13, 1.84)
Male	301	57	Reference group
D4Z4 repeat length			
1–3	60	24	4.14 (2.87, 5.67)
4–7	340	76	Reference group
8–10	119	17	0.56 (0.40, 0.78)

entry were predicted to have a higher likelihood of wheelchair use. Genetics (all repeat lengths) and presenting symptoms were lower down on the list of features predicted to influence wheelchair use.

Medication-only model

A total of 494 participants reported taking 1461 unique medications that were grouped into 359 medication categories. Patients not taking medications ($n = 84$) were excluded from this model. As expected, the accuracy was lower at 0.62 and the AUC was 0.66. All classes of medications were predicted to increase risk of wheelchair use with the exception being those that were classified as amino acids (Fig. 4C).

Discussion

This study is one of the largest longitudinal datasets described to date in FSHD and supports observations from mostly cross-sectional studies, and also provides some potential new areas of insights to improve patient care. Here we showed a relationship between repeat length, age at diagnosis, and age at first wheelchair use, with a higher risk of wheelchair use overall in females. When considering all the clinical data collected, machine learning analysis suggested that the feature most predictive of wheelchair use was disease duration, with genetics playing a smaller role, and that medical comorbidities may also impact motor outcomes in FSHD.

Our data supports previous observations that individuals with smaller allele sizes (1–3 D4Z4 repeat units) are diagnosed at a younger age and are more likely to use a wheelchair at a younger age compared to those with larger allele sizes.^{9,12,23,24} However, the machine learning analysis showed that longer disease duration had a larger influence on progression to wheelchair use than genetics. While genetic mutation clearly has an impact on the age of diagnosis and motor outcomes, over the course of a clinical trial (~1 year) the predictive value is much less clear. It may be that stratifying trial participants by disease duration may provide more useful trial planning information.

Traditionally, males with FSHD are thought to be more severely affected than females, and that females tend to be diagnosed at an older age compared to males.^{11,25–27} It has been hypothesized that hormonal differences may play a role in the different clinical outcomes observed between males and females. Ricci et al.²⁵ showed that males with FSHD and their male relatives developed symptoms of motor impairment earlier than females, an observation that was first observed around age 20 and ended around age 50. Subsequent *in vitro* studies using myoblast cell cultures from patients with FSHD showed that exposure to oestrogen improved myoblast cell differentiation through decreased transcriptional activity of the double homeobox 4 (DUX4) protein and interference with recruitment of DUX4 in the nucleus, suggesting a protective effect of oestrogens.²⁸ Aberrant expression of DUX4 is felt to be one of the pathogenic mechanisms contributing to decreased

myoblast cell survival in patients with FSHD. Mul et al.²⁹ evaluated the lifetime endogenous exposure to oestrogen and found no significant effect, protective or damaging, to account for the reported clinical variability seen between males and females. Banerji et al.⁵ recently reported that females who had been pregnant or carried multiple children to term were associated with a slower onset of muscle weakness. Contrary to this, Ciafalonni et al.³⁰ found that 24% of females reported worsening of their FSHD symptoms following childbirth that did not improve.

In our study, we found that females were diagnosed approximately 10 years later than males in the 4–7 D4Z4 repeat category. The separation in age at diagnosis appeared around age 20 and disappeared around age 50, ages that roughly correlate with onset of puberty and menopause, respectively. On the other hand, we found a higher frequency of females in the 1–3 D4Z4 repeat group who were diagnosed on average 5 years earlier than males, although this was not statistically significant. Even after adjusting for differences in allele length, we found that females were more likely to progress to wheelchair use and at a faster rate compared to males. While this does not explain the difference in age at diagnosis based on genetics, it does raise the possibility that females are more severely affected by the disease than previously thought. One possible explanation is that females have a more insidious disease onset as compared to males leading to older age at diagnosis, or that their concerns are not taken seriously by physicians, leading to the observed older age at diagnosis yet faster progression to wheelchair use. It may also be the case that females are more likely than males to use assistive devices, which may make it appear as though they have faster progression to wheelchair use. Registry data include information on whether or not participants have children, and if those children are affected by FSHD, but does not include information on disease progression following childbirth for females, although this may be worthwhile to include in the future given contradictory reports^{5,30} of disease progression following childbearing. Additionally, large scale studies will be needed to further clarify these findings and better understand if there are sex and/or hormonal differences that contribute to clinical presentation and functional outcomes.

This study used machine learning technology to evaluate all clinical data collected during enrolment in The Registry and on the annual follow-up questionnaires, providing a unique opportunity for longitudinal analysis of individual disease progression over time. The model identified several age-related items high on the list of features influencing wheelchair use. In addition, the model identified medical comorbidities and medication use as important features influencing wheelchair use. To our knowledge, this is the first time that medical comorbidities and medication use have been associated with functional outcomes (wheelchair use) in patients with FSHD. Surprisingly, genetics and presenting symptoms were lower down on the list of features identified as having an influence on wheelchair use.

The model identified breathing difficulty as the medical comorbidity having the most influence on progression to wheelchair use. FSHD can result in restrictive lung disease and studies have shown that expiratory muscles, rather than diaphragmatic muscles, tend to be more affected.^{31–33} Santos et al.³² evaluated 29 age- and sex-matched patients with FSHD with and without respiratory dysfunction and found that patients with respiratory dysfunction had involvement of expiratory musculature and 20/29 met criteria to start (non-invasive) mechanical ventilation, 14 of whom were wheelchair bound. It is not clear from our study if breathing difficulty as a consequence of FSHD is truly predictive of wheelchair use, or if this is a medical condition that suggests worsening disease status and is therefore seen in higher frequency in individuals using a wheelchair. Future studies will be needed to better

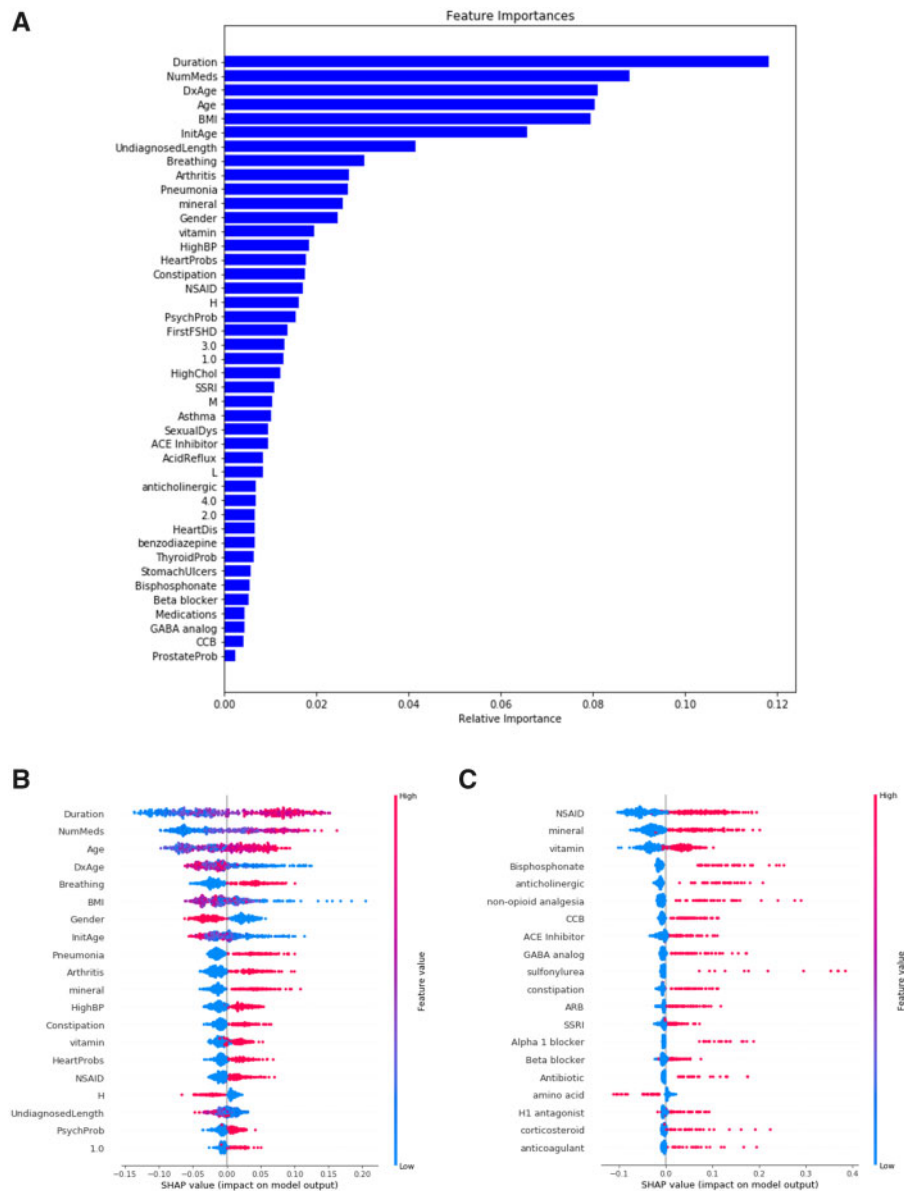


Figure 4 Feature importance predicted by the Random Forest machine learning model and SHAP analysis. The Random Forest machine learning model (A) and SHAP analysis (B) both identified disease duration and number of medications as the most important features influencing wheelchair use. Age-related features such as current age of the patient (Age), age at diagnosis (DxAge) were the next most important features. Female gender was found to influence likelihood of wheelchair use. Having a low BMI was found to influence towards wheelchair use. Comorbidities such as respiratory concerns (Breathing), arthritis, pneumonia, hypertension (HighBP) and constipation were all found to influence towards wheelchair use. Genetics (repeat length) and initial presenting symptoms were further down on the list of feature importance. A separate ‘medication-only’ model (C) found that all classes of medications influenced towards wheelchair use except for those classified as amino acids. Duration = disease duration; NumMeds = number of medications; Age = current age of the patient; DxAge = age at diagnosis; Breathing = Y/N respiratory difficulties; InitAge = initial age symptom onset; HighBP = hypertension; HeartProbs = heart problems; H = 8–10 D4Z4 repeat units; UndiagnosedLength = time spent undiagnosed (in years); PsychProb = psychiatric concerns; 1.0 = initial symptom facial weakness

understand how breathing difficulty relates to disease progression and functional outcomes in FSHD, and if this can be used as a way to monitor disease progression over time.

Additional medical comorbidities identified as adding to risk of wheelchair use included arthritis, pneumonia, hypertension, constipation and psychological problems (such as depression or anxiety); however, not having these problems was not predicted to influence away from wheelchair use. All of the comorbidities identified as influencing wheelchair use can be treated with various medications. The model predicted that the more medications one takes, the higher the likelihood of using a wheelchair. This raises several interesting questions about medical management of

comorbidities in patients with FSHD and their overall level of general health as it relates to risk of wheelchair use. Research suggests that ~50% of patients do not take their medications as prescribed, leading to increased morbidity and mortality.³⁴ Fitzgerald et al.³⁵ surveyed participants in The Registry to evaluate medication adherence in patients with FSHD and myotonic dystrophy and found that 44% of patients with FSHD had hypertension, followed by arthritis (29.5%) and depression (28.5%). One-third of participants with FSHD reported taking more than six medications daily (prescription and over-the-counter), and those individuals were more likely to be older and unemployed compared to those taking fewer medications. Most participants (82.2%) reported good medication

compliance without significant barriers (defined as cost of medication, side effects and understanding of need for medication) to taking their medications as prescribed. These findings suggest that there may be an association between the number of medications a patient is taking, medical comorbidities and progression to wheelchair use. Future studies will be needed to determine if better medical management of comorbidities might influence functional outcomes in FSHD, or if loss of mobility due to FSHD predisposes individuals to developing comorbidities.

We further investigated the different classes of medications reported by The Registry participants to see if there were specific types of medications that influenced wheelchair use. The medication-only model is inherently less accurate but did show that all types of medications influenced towards wheelchair use except those classified as amino acids, which included supplements such as acetyl-L-carnitine, L-lysine, branched chain amino acids, N-acetyl cysteine and hydroxymethylbutyrate (data not shown). Interestingly, the general model that included all reported clinical data identified minerals, vitamins and non-steroidal anti-inflammatory drugs (NSAIDs) as having the highest influence on wheelchair use. Here we cannot determine whether the number of medications and medical comorbidities increased the likelihood of needing a wheelchair, or whether more severely affected individuals were more likely to have medical comorbidities necessitating treatment. These findings will require a more thorough investigation using larger datasets to better understand the influence of medication use and medical comorbidities on patient outcomes in FSHD.

There are several limitations to this study. The Registry is a collection of patient-reported symptoms and medical data that are collected annually, raising the possibility of recall bias. In an effort to minimize this, we focused on features that are memorable moments in one's lifetime such as age at diagnosis or age at first wheelchair use. There is also the possibility of selection bias towards those individuals willing to participate in The Registry. A large proportion of individuals enrolled in The Registry were excluded from this analysis because they were not genetically defined, which could result in further selection bias. We chose to focus only on individuals who are genetically confirmed to have FSHD1 to minimize the possibility that individuals are included for analysis who do not truly have FSHD and improve the reliability of our results for those with FSHD1. The Registry contains only a small number of individuals with FSHD2, and it is possible that some of the individuals who were not genetically defined have FSHD2. Although these individuals are felt to have a clinical course similar to those with FSHD1, they were excluded from this analysis in an effort to minimize confounding variables. Individuals in the registry were primarily Caucasian and future enrolment should aim to improve diversity and include more individuals from different ethnic backgrounds to better represent the spectrum of disease in the population. We also found that most individuals participating in The Registry were highly educated, raising the possibility that more educated individuals are more likely to participate in The Registry. Almost half of individuals were unemployed at the time of enrolment. This could represent registry bias towards individuals who are more severely affected, whereas individuals who are still working and not as severely affected may be less inclined to join. Finally, from a machine learning perspective, the number of subjects analysed is on the smaller side of what would be traditionally used to build a model. Future studies should aim to increase the size of the dataset to provide further validation of the findings reported here.

In conclusion, we showed that while genetics and gender may influence age at diagnosis, machine learning technology suggests that medical comorbidities and medication use may have a larger

influence on functional outcomes for patients with FSHD than previously appreciated. We found an association between the number of medications one takes and medical comorbidities, but the direction of this relationship is not clear. Future studies should aim to clarify this relationship to help determine if aggressive medical management of comorbidities can improve functional outcomes in patients with FSHD. This could also have future implications for clinical trial design by restructuring how patients are categorized (e.g. by disease duration) and identifying clinically meaningful outcome measures to assess for improvement in function over time.

Acknowledgements

We would like to thank the patients and families who have chosen to participate in The Registry and made this work possible. We would like to thank the FSHD Society for their generous support to make this project possible. We would also like to thank Bill Martens from The Registry for his help with data acquisition and clarification. We would also like to thank Dr Jean-Baptiste Le Pichon, MD for his guidance and feedback throughout the project development and manuscript preparation.

Funding

This work was supported by a grant from the FSHD Society (FSHD 2020—SG01: AI Proof-of-Concept for FSHD Research). The Registry is supported by the National Institutes of Health and National Institute of Arthritis and Musculoskeletal and Skin Diseases (contracts #N01-AR-5-2274 and #N01-AR-0-2250), and the Senator Paul D. Wellstone Muscular Dystrophy Cooperative Research Centers (grant #U54-NS048843).

Competing interests

J.S. reports grant funding from NIH, MDA, FSHD Society, and Friends of FSH Research. J.S. is a consultant or serves on the advisory board for Dyne, Avidity, Fulcrum, Acceleron, MT Pharma, VectivBio, and Sarepta.

References

1. Flanigan KM, Coffeen CM, Sexton L, Stauffer D, Brunner S, Leppert MF. Genetic characterization of a large, historically significant Utah kindred with facioscapulohumeral dystrophy. *Neuromuscul Disord.* 2001;11(6-7):525–529.
2. Mostacciolo ML, Pastorello E, Vazza G, et al. Facioscapulohumeral muscular dystrophy: Epidemiological and molecular study in a north-east Italian population sample. *Clin Genet.* 2009;75(6):550–555.
3. Padberg GW, Frants RR, Brouwer OF, Wijmenga C, Bakker E, Sandkuijl LA. Facioscapulohumeral muscular dystrophy in the Dutch population. *Muscle Nerve Suppl.* 1995;2(2):S81–S84.
4. Deenen JC, Arnts H, van der Maarel SM, et al. Population-based incidence and prevalence of facioscapulohumeral dystrophy. *Neurology.* 2014;83(12):1056–1059.
5. Banerji CRS, Cammish P, Evangelista T, Zammit PS, Straub V, Marini-Bettolo C. Facioscapulohumeral muscular dystrophy 1 patients participating in the UK FSHD registry can be subdivided into 4 patterns of self-reported symptoms. *Neuromuscul Disord.* 2020;30(4):315–328.
6. Statland JM, Tawil R. Facioscapulohumeral muscular dystrophy. *Continuum (Minneapolis Minn).* 2016;22(6, Muscle and Neuromuscular Junction Disorders):1916–1931.

7. Preston MK, Tawil R, Wang LH, et al. Facioscapulohumeral muscular dystrophy. In: Adam MP, Ardinger HH, Pagon RA, eds. *GeneReviews*(®) [Internet]. University of Washington; 1993.
8. Tawil R. Facioscapulohumeral muscular dystrophy. *Handb Clin Neurol*. 2018;148:541–548.
9. Goselink RJM, Mul K, van Kernebeek CR, et al. Early onset as a marker for disease severity in facioscapulohumeral muscular dystrophy. *Neurology*. 2019;92(4):e378–e385.
10. Goselink RJM, van Kernebeek CR, Mul K, et al. A 22-year follow-up reveals a variable disease severity in early-onset facioscapulohumeral dystrophy. *Eur J Paediatr Neurol*. 2018;22(5):782–785.
11. Statland JM, Donlin-Smith CM, Tapscott SJ, Lemmers RJ, van der Maarel SM, Tawil R. Milder phenotype in facioscapulohumeral dystrophy with 7-10 residual D4Z4 repeats. *Neurology*. 2015;85(24):2147–2150.
12. Statland JM, Tawil R. Risk of functional impairment in facioscapulohumeral muscular dystrophy. *Muscle Nerve*. 2014;49(4):520–527.
13. Lamperti C, Fabbri G, Vercelli L, et al. A standardized clinical evaluation of patients affected by facioscapulohumeral muscular dystrophy: The FSHD clinical score. *Muscle Nerve*. 2010;42(2):213–217.
14. Andersen G, Dahlqvist JR, Vissing CR, Heje K, Thomsen C, Vissing J. MRI as outcome measure in facioscapulohumeral muscular dystrophy: 1-year follow-up of 45 patients. *J Neurol*. 2017;264(3):438–447.
15. Ferguson MR, Poliachik SL, Budech CB, et al. MRI change metrics of facioscapulohumeral muscular dystrophy: Stir and T1. *Muscle Nerve*. 2018;57(6):905–912.
16. Sun Z, Ghosh S, Li Y, et al. A probabilistic disease progression modeling approach and its application to integrated Huntington's disease observational data. *JAMIA Open*. 2019;2(1):123–130.
17. Handelman GS, Kok HK, Chandra RV, Razavi AH, Lee MJ, Asadi H. eDoctor: machine learning and the future of medicine. *J Intern Med*. 2018;284(6):603–619.
18. Kavakiotis I, Tsave O, Salifoglou A, Maglaveras N, Vlahavas I, Chouvarda I. Machine learning and data mining methods in diabetes research. *Comput Struct Biotechnol J*. 2017;15:104–116.
19. Zou Q, Qu K, Luo Y, Yin D, Ju Y, Tang H. Predicting diabetes mellitus with machine learning techniques. *Front Genet*. 2018;9:515.
20. Tseng PY, Chen YT, Wang CH, et al. Prediction of the development of acute kidney injury following cardiac surgery by machine learning. *Crit Care*. 2020;24(1):478.
21. Sun YV. Multigenic modeling of complex disease by random forests. *Adv Genet*. 2010;72:73–99.
22. Lundberg SM, Lee SI. A unified approach to interpreting model predictions. In: *31st Conference on Neural Information Processing Systems (NIPS)*; 2017:1–10.
23. Lunt PW, Jardine PE, Koch MC, et al. Correlation between fragment size at D4F104S1 and age at onset or at wheelchair use, with a possible generational effect, accounts for much phenotypic variation in 4q35-facioscapulohumeral muscular dystrophy (FSHD). *Hum Mol Genet*. 1995;4(5):951–958.
24. Ricci E, Galluzzi G, Deidda G, et al. Progress in the molecular diagnosis of facioscapulohumeral muscular dystrophy and correlation between the number of KpnI repeats at the 4q35 locus and clinical phenotype. *Ann Neurol*. 1999;45(6):751–757.
25. Ricci G, Scionti I, Sera F, et al. Large scale genotype-phenotype analyses indicate that novel prognostic tools are required for families with facioscapulohumeral muscular dystrophy. *Brain*. 2013;136(Pt 11):3408–3417.
26. Tonini MM, Passos-Bueno MR, Cerqueira A, Matioli SR, Pavanello R, Zatz M. Asymptomatic carriers and gender differences in facioscapulohumeral muscular dystrophy (FSHD). *Neuromuscul Disord*. 2004;14(1):33–38.
27. Zatz M, Marie SK, Cerqueira A, Vainzof M, Pavanello RC, Passos-Bueno MR. The facioscapulohumeral muscular dystrophy (FSHD1) gene affects males more severely and more frequently than females. *Am J Med Genet*. 1998;77(2):155–161.
28. Teveroni E, Pellegrino M, Sacconi S, et al. Estrogens enhance myoblast differentiation in facioscapulohumeral muscular dystrophy by antagonizing DUX4 activity. *J Clin Invest*. 2017;127(4):1531–1545.
29. Mul K, Horlings CGC, Voermans NC, Schreuder THA, van Engelen BGM. Lifetime endogenous estrogen exposure and disease severity in female patients with facioscapulohumeral muscular dystrophy. *Neuromuscul Disord*. 2018;28(6):508–511.
30. Ciafaloni E, Pressman EK, Loi AM, et al. Pregnancy and birth outcomes in females with facioscapulohumeral muscular dystrophy. *Neurology*. 2006;67(10):1887–1889.
31. D'Angelo MG, Romei M, Lo Mauro A, et al. Respiratory pattern in an adult population of dystrophic patients. *J Neurol Sci*. 2011;306(1-2):54–61.
32. Santos DB, Boussaid G, Stojkovic T, et al. Respiratory muscle dysfunction in facioscapulohumeral muscular dystrophy. *Neuromuscul Disord*. 2015;25(8):632–639.
33. Stubgen JP, Schultz C. Lung and respiratory muscle function in facioscapulohumeral muscular dystrophy. *Muscle Nerve*. 2009;39(6):729–734.
34. Brown MT, Bussell JK. Medication adherence: WHO cares? *Mayo Clinic Proc*. 2011;86(4):304–314.
35. Fitzgerald BP, Conn KM, Smith J, et al. Medication adherence in patients with myotonic dystrophy and facioscapulohumeral muscular dystrophy. *J Neurol*. 2016;263(12):2528–2537.