

Gait Characteristics and Their Discriminative Ability in Patients with Fabry Disease with and Without White-Matter Lesions

José Braga[1], Flora Ferreira[1,2(✉)], Carlos Fernandes[1], Miguel F. Gago[3,4], Olga Azevedo[3,4], Nuno Sousa[4], Wolfram Erlhagen[2], and Estela Bicho[1]

[1] Center Algoritmi, University of Minho, Guimarães, Portugal
estela.bicho@dei.uminho.pt
[2] Center of Mathematics, University of Minho, Guimarães, Portugal
{fjferreira,wolfram.erlhagen}@math.uminho.pt
[3] Neurology and Cardiology Services,
Hospital Senhora da Oliveira, Guimarães, Portugal
miguelfgago@yahoo.com
[4] ICVS, School of Medicine, University of Minho, Braga, Portugal

Abstract. Fabry disease (FD) is a rare disease commonly complicated with white matter lesions (WMLs). WMLs, which have extensively been associated with gait impairment, justify further investigation of its implication in FD. This study aims to identify a set of gait characteristics to discriminate FD patients with/without WMLs and healthy controls. Seventy-six subjects walked through a predefined circuit using gait sensors that continuously acquired different stride features. Data were normalized using multiple regression normalization taking into account the subject physical properties, with the assessment of 32 kinematic gait variables. A filter method (Mann Whitney U test and Pearson correlation) followed by a wrapper method (recursive feature elimination (RFE) for Logistic Regression (LR) and Support Vector Machine (SVM) and information gain for Random Forest (RF)) were used for feature selection. Then, five different classifiers (LR, SVM Linear and RBF kernel, RF, and K-Nearest Neighbors (KNN)) based on different selected set features were evaluated. For FD patients with WMLs versus controls the highest accuracy of 72% was obtained using LR based on 3 gait variables: pushing, foot flat, and maximum toe clearance 2. For FD patients without WMLs versus controls, the best performance was observed using LR and SVM RBF kernel based on loading, foot flat, minimum toe clearance, stride length variability, loading variability, and lift-off angle variability with an accuracy of 83%. These findings are the first step to demonstrate the potential of machine learning techniques based on gait variables as a complementary tool to understand the role of WMLs in the gait impairment of FD.

This work has been supported by FCT – Fundação para a Ciência e Tecnologia within the R&D Units Project Scope: UIDB/00319/2020 and UIDB/00013/2020.

© Springer Nature Switzerland AG 2020
O. Gervasi et al. (Eds.): ICCSA 2020, LNCS 12251, pp. 415–428, 2020.
https://doi.org/10.1007/978-3-030-58808-3_30

Keywords: Fabry Disease · Feature selection · Machine learning · Gait

1 Introduction

Fabry Disease (FD) is a rare genetic X-linked lysosomal disorder caused by the deficiency or absent activity of the enzyme α-galactosidase A, resulting in the accumulation of globotriaosylceramide (GL-3) in several organs, including the kidney, heart and brain [10]. Brain manifestations in FD include progressive white matter lesions (WMLs) [4,14]. Brain WMLs were an early manifestation, affecting 11.1% of males and 26.9% of females under 30 years of age, even without cerebrovascular risk factor outside FD [3]. In [14], WMLs were found in 46% of 1276 patients which tend to occur earlier in males and their prevalence revealed to increase with patients' age. White matter is the brain region responsible for the transmission of nerve signals and for communication between different parts of the brain. WMLs have been associated with gait impairment [23] and the risk of falls [22,27]. Gait abnormalities, such as slower gait and postural instability, have been reported in FD [16]. However, not much is known regarding the impact of WMLs on gait performance in patients with FD. Specific location and distribution of WMLs suggest a specific underlying disease [14] and may reflate in a different gait profile. In fact, gait evaluation can be useful to differentiate different pathologies even in the presence of highly overlapping phenotypes, such as the differences found between two types of Parkinsonism (Vascular Parkinsonism versus Idiopathic Parkinson's Disease) in [9]. Furthermore, gait assessment has been revealing as a good complementary clinical tool to discriminate adults with and without a pathology such as Parkinson's disease [8,9,15], Huntington's disease [17] and recently FD [7].

Gait is usually described by its spatio-temporal and foot clearance characteristics such as speed, stride length, stride time, minimum toe clearance (mean gait characteristics), and their respective variability (given by the coefficient of variation or the standard deviation) [9,15,21]. Different machine learning (ML) techniques have been used to select the best combination of relevant gait characteristics for gait classification [6,8,21]. Recent work [21] shows that a subset of gait characteristics selected using random forest with information gain and recursive features elimination (RFE) technique with Support Vector Machine (SVM) and Logistic Regression improves the classification accuracy of Parkinson's Disease. Widely used machine learning models for gait classification are SVM [1,7,17,20,26], Random Forest (RF) [1,7,26] and K-Nearest Neighbor (KNN) [20,26]. The outcomes of these studies show good performance accuracy in the classification of pathological gait. In particular, the outcomes of previous work [7] show promising results in the use of gait characteristics to discriminate FD patients and healthy adults. However, the implication of the presence or absence of WMLs in the gait performance has not yet been investigated. Hence, the purpose of this study is to identify relevant gait characteristics to discriminate FD patients with and without WMLs from healthy subjects (controls).

Our hypothesis is that from the evaluation of different gait characteristics (e.g., stride length, stride time variability) one can (1) identify unique gait characteristics in the two groups of FD and (2) a selection of relevant gait characteristics can accurately discriminate FD patients with WMLs versus healthy adults, and FD patients without WMLs versus healthy adults.

2 Materials and Methods

2.1 Experimental Protocol

Data from 39 patients with FD (25 with WMLs and 14 without WMLs) and 37 healthy controls were collected. From the control group was constructed two groups of controls aged-matched with the groups of FD patients with WMLs and FD patients without WMLs. The subject demographics of these four groups are summarized in Table 1. For all FD patients, the exclusion criteria were: less than eighteen years of age, the presence of resting tremor, moderate-severe dementia (CDR $>$ 2), depression, extensive intracranial lesions or neurodegenerative disorders, musculoskeletal disease, and rheumatological disorders. Local hospital ethics committee approved the protocol of the study, submitted by ICVS/UM and Center Algoritmi/UM. Written consent was obtained from all subjects or their guardians.

Table 1. Demographic variables for FD patients with and without WMLs and Controls.

Demographics	Fabry with WMLs (n = 25)	Controls (n = 25)	p-value	Fabry without WMLs (n = 14)	Controls (n = 14)	p-value
Age (years)	57.92 (16.75)	60.52 (22.04)	0.174*	37.79 (10.48)	38.07 (12.57)	0.946*
Male (%)	32%	32%	1†	43%	43%	1†
Weight (kg)	65.94 (10.05)	64.68 (8.46)	.648*	64.75 (6.67)	68.89 (13.25)	.482*
Height (m)	1.59 (0.07)	1.63 (0.06)	.052*	1.66 (0.09)	1.69 (0.08)	.246*

Data is presented as mean (standard deviation). *Mann-Whitney U test. †Fisher Exact T Test

Two Physilog® sensors (Gait Up®, Switzerland) positioned on both feet were used to measure different gait variables of each stride. This study consists of walking a 60-m continuous course (30 m corridor with one turn) in a self-selected walking while the sensors are acquiring the data. This data consists in the arithmetic mean calculated for all subjects stride time series for 11 spatial-temporal variables: speed (velocity of forward walking for one cycle), cycle duration (duration of one cycle), cadence (number of cycles in a minute), stride length (distance between successive footprints on the ground for the same foot), stance (time in which the foot touches the ground), swing (time in which the foot is in the air and does not touch the ground), loading (percent of stance between the heel strike and the foot being entirely on the ground), foot flat (percent of stance where the foot is entirely on the ground), pushing (percent of stance between the foot being entirely on the ground and toe leaving the ground), double support (percent of the cycle where both feet are touching the ground), peak swing

(maximum angular velocity during swing) and 6 foot clearance variables strike angle (angle between the foot and ground when the heel touches the ground), lift-off angle (angle between the foot and the ground at take-off), maximum heel clearance (maximum height reached by the heel), maximum toe clearance 1 (maximum height reached by the toes just after maximum heel clearance), minimum toe clearance (minimum height of the toes during swing) and maximum toe clearance 2 (maximum height reached by the toes just before heel strike) in a total of 17 gait variables that contain the full step data. The average walking speed of FD patients and controls was 1.335 ± 0.174 m/s and 1.333 ± 0.201 m/s, respectively.

2.2 Data Normalization

Gait characteristics of a subject are affected by his demographics properties including height, weight, age, and gender, as well as by walking speed [25,26] or stride length [2,7,9]. To normalize the data regression models according to Wahid et al.'s method [25,26] were used. Comparing to other methods, such as dimensionless equations and detrending methods, MR normalization revealed better results on reducing the interference of subject-specific physical characteristics and gait variables [18,25,26], thereby improving gait classification accuracy using machine learning methods [7,26].

First, to control the multicollinearity among predictor variables within this multiple regression, Variance Inflation Factor (VIF) was calculated [24]. This test measures the colinearity of the physical characteristics (age, weight, height, and gender), speed, and stride length, being a value of VIF greater than 5 an indicator of this strong correlation. In Table 2 are the results of VIF tests. Since, when testing VIF with all the variables, both Speed and Stride Length had a VIF higher than 5 (an indicator of strong correlation), they can not be used simultaneously. When used separately all VIF values are less than 5, so there is no evidence of multicollinearity.

Table 2. Variance inflation factor for physical characteristics (age, weight, height and gender), speed and stride length.

	Age	Weight	Height	Gender	Speed	Stride length
VIF	2.24	1.97	3.76	1.28	6.05	9.89
	1.99	1.93	3.37	1.27	–	2.06
	1.82	1.97	2.9	1.28	1.26	–

Spatio-temporal gait variables and foot clearance variables were normalized as follows:

$$\hat{y}_i = \beta_0 + \sum_{j=1}^{p} \beta_j x_{ij} + \varepsilon_i \tag{1}$$

where $\hat{y}_i$ represents the prediction of the dependent variable for the ith observation; x_{ij} represents the jth physical property of the ith observation including age, weight, height, gender, speed or stride length, β_0 represents the intercept term, β_j represents the coefficient for the jth physical property and ε_i represents the residual error for the ith observation. The model's coefficients are estimated using the physical properties and the mean values of the gait variables of the 37 healthy controls. Although at least 20 subjects per independent variable are recommended in multiple linear regression [13], based on similar studies [18,25] with higher sample sizes, MR models were computed for all combinations with 1, 2, and 3 independent variables using a bisquare weight function. For the models with all significant independent variables ($p < 0.05$) Akaike's information criterion (AIC) [5] and R-squared metrics were used to select the best-fitted model. Statistical assumptions of a linear regression including linearity, normality, and homoscedasticity were verified.

The models created for each feature are summarized in Table 3. These are similar to the ones found in [7] with 34 controls. Left foot models were also computed but no major differences were found among both feet. So, in this paper, we just report the results on the right foot.

Table 3. Multiple linear regression models using only significative independent variables for each gait variable on the right foot

	MR normalization							AIC	Adjusted R^2
		Age	Weight	Height	Speed	Gender	Stride length		
Spatial-temporal variables									
Cycle duration	= 1.083	-0.0011		0.218	-0.274			-124.52	0.658
Cadence	= 110.429	0.133		-24.805	31.282			220.24	0.697
Stride length	= 0.255	-0.0019		0.310	0.506			-114.77	0.889
Stance	= 65.538				-4.291	-0.362		156.56	0.130
Swing	= 34.462				4.291	0.362		156.56	0.130
Loading	= 17.464		-0.131		2.214	2.427		157.45	0.371
Foot flat	= 59.319	0.0851	0.166		-15.311			194.08	0.677
Pushing	= 24.281	-0.075	-0.076		13.336			186.14	0.629
Double support	= 32				-8.804	-1.454		197.81	0.207
Peakswing	= 284.564			-68.139	162.419	16.1879		344.59	0.637
Foot clearance variables									
Strike angle	= 17.135		-0.196			4.811	14.806	184.33	0.640
Lift-off angle	= -41.883	0.209					-27.099	219.29	0.770
Maximum heel clearance	= 0.219	-0.0006				0.039	0.039	-150.06	0.435
Maximum toe clearance 1	= 0.056	0.00021				0.0085		-171.88	0.018
Minimum toe clearance	= 0.012	0.00032						-248.47	0.441
Maximum toe clearance 2	= 0.060	-0.00061				0.027	0.088	-183.44	0.711

In each subject group, the best fitted MR models are used to normalize each stride gait variable by dividing the original value y_i by the predicted gait variable $\hat{y}_i$ from (1), as follow:

$$y_i^n = \frac{y_i}{\hat{y}_i} \tag{2}$$

where y_i^n represents the normalized value for the ith observation.

After normalizing all strides of each of the 16 gait variables, the mean and the standard deviation (SD) of each variable (each gait time series) for all subjects were calculated. In this work, the SD value is used to measure the variability of each gait variable.

2.3 Feature Selection

Due to the high number of gait variables (a total of 32) and a small number of samples a hybrid method (filter method followed by a wrapper method) was employed to select the most relevant gait characteristics. First, a filter method based on Mann Whitney U tests and Spearman's correlation between the variables was used to selected 12 gait characteristics. Mann Whitney U Tests were used to examine the difference between groups (FD with WMLs vs. Controls and FD without WMLs vs. Controls) for each gait characteristics and Spearman's correlation to evaluate the independence and redundancy between them. Therefore, the selected 12 gait characteristics are the ones that present higher U-value and do not present a high correlation between them ($\rho < |0.9|$). Before applying a wrapper method, the selected 12 gait variables were scaled to have zero mean and unit variance. Based on previous work [21], the wrapper was developed using the Recursive Feature Elimination (RFE) technique with three different ML classifiers: Logistic Regression (LR), SVM with Linear kernel and RF. RFE has some advantages over other filter methods [12]. RFE is an iterative method where features are removed one by one without affecting the training error. The selection of the optimal number of features for each model is based on the evaluation metric F1 Score evaluated through 5-fold cross-validation. F1 score is defined as the harmonic mean between precision and recall whose range goes from 0 to 1 and tells how precise the classifier is [11]. The gait characteristics' importance was quantified using the model itself (feature importance for LR and SVM with linear kernel and information gain for RF). The F1 score was used to assess the performance of the different gait characteristics combinations.

2.4 Machine Learning Classifiers

Based on the literature [1,7,17,20,21,26] four different types of classifiers were employed for the purpose of distinguishing FD patients with and without WMLs versus controls: LR, SVM, RF and KNN.

All classifiers hyperparameters were tuned using randomized search and grid search method with 5-fold cross-validation: LR regularization strength constraint and type of penalty according to regularization strength constraint; SVM

regularization parameter, gamma and different types of kernel (with different degrees); RF maximum depth, minimum samples leaf, minimum samples split and the number of estimators; finally, KNN number of neighbors, weights and metric. All classifiers were implemented in Python programming language using Scikit-learn library [19].

To evaluate the performance of the different classification models the accuracy metric (the ratio of correct predictions) of the training and validation set was used.

3 Results

3.1 FD Patients with WMLs Versus Controls

Recursive feature elimination technique was performed on the 12 remaining features from filter method: cycle duration, cadence, swing, foot flat, pushing, double support, maximum heel clearance, maximum toe clearance 1, maximum toe clearance 2, loading variability, foot flat variability, and maximum heel clearance variability. The results are summarized in Table 4 and Fig. 1. Five gait characteristics were selected by LR with an F1 Score of 63.96%, while SVM selected 4 features with an F1 score of 63.96%, and RF selected the larger number of features, 9 with an F1 score of 60.21%. In Table 4, are the training and validation accuracies for the optimal models of each algorithm. RF had the higher accuracy in the training but the lower accuracy in validation. LR and SVM show similar accuracies in training and validation.

Table 4. Classification model accuracy for FD patients with WMLs versus controls

	Training accuracy % (Mean ± SD)	Validation accuracy % (Mean ± SD)
Logistic Regression	70.50 ± 4.85	70 ± 14.14
SVM (Linear Kernel)	74.50 ± 3.32	70 ± 10.95
Random Forest	86 ± 2.55	66 ± 13.56

Figure 1 (right side) shows the contribution of each gait characteristic in the classification model. The common gait characteristics among the selected features by each model were three (Top 3): foot flat, pushing, and maximum toe clearance 2. The common gait characteristics from LR and SVM were four (Top 4): foot flat, pushing, maximum toe clearance 2, and foot flat variability. These Top 3 and Top 4 were evaluated with five classification models (LR, SVM Linear kernel, SVM RBF kernel, RF, and KNN) to identify the optimal combination of gait characteristics and the classification model with better performance. Results are presented in Table 5. LR performance increases slightly by

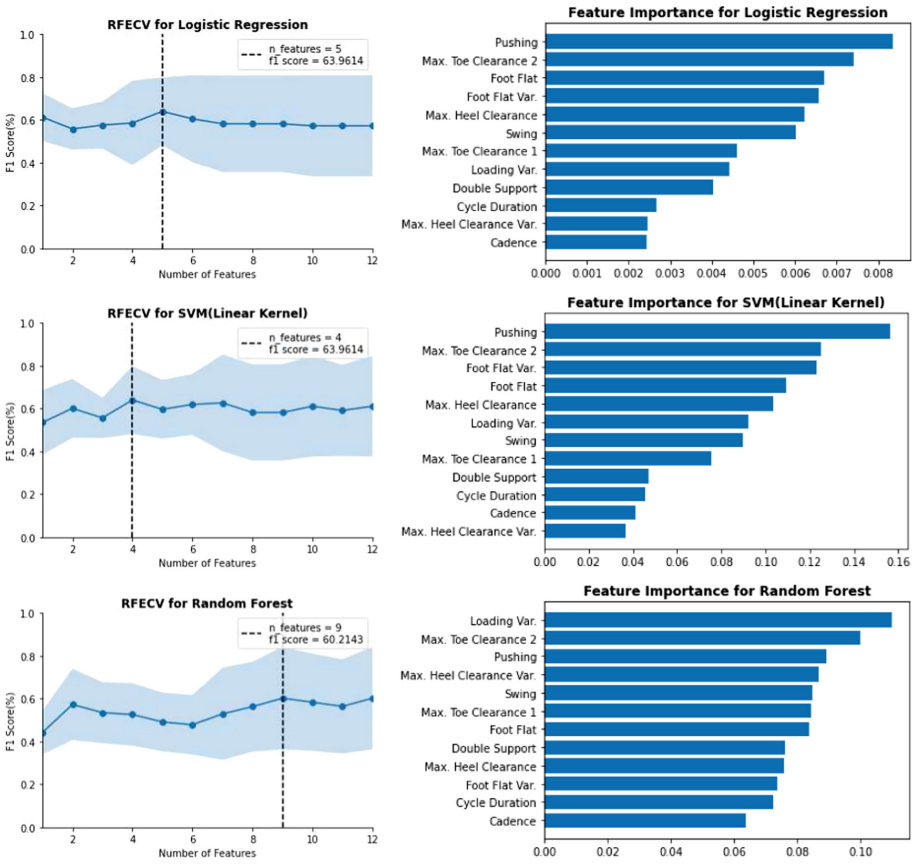

Fig. 1. F1 score versus number of features for selection of optimal numbers of gait characteristics (left) and feature importance results (right) obtained based on Logistic Regression, SVM Linear kernel and Random Forest in FD patients with WMLs vs. controls. Recursive features elimination was used through the 5-fold cross-validation (RFECV). Max.: maximum; Var.: variability.

reducing the number of features to 3. RF shows similar training and validation results with the selected features by correspondent ML model (Table 4), Top 3, and Top 4 (Table 5). With Top 4, higher accuracies are observed in SVM Linear kernel and SVM RBF kernel related to Top 3. In fact, the Top 4 corresponds to the features selected with SVM Linear Kernel. KNN shows equal validation accuracy with Top 3 and Top 4, although it displays higher training accuracy with Top 4. The better performance is observed for LR with Top 3 achieving a validation accuracy of 72% and a training accuracy of 71%, followed by both SVM Linear kernel and SVM RBF kernel with Top 4 achieving around 70% validation accuracy and 74% training accuracy.

Table 5. Classification model accuracy for top common gait characteristics in FD patients with WMLs versus controls

	Top 3 common features (LR, SVM and RF)		Top 4 common features (LR and SVM)	
	Training accuracy (Mean ± SD)	Validation accuracy (Mean ± SD)	Training accuracy (Mean ± SD)	Validation accuracy (Mean ± SD)
LR	71 ± 3.74	72 ± 13.27	74 ± 4.64	70 ± 14.14
SVM (Linear Kernel)	71.50 ± 4.90	66 ± 14.97	74.50 ± 3.32	70 ± 10.95
SVM (RBF kernel)	72 ± 4.85	66 ± 14.97	74 ± 3.39	70 ± 10.95
RF	85 ± 2.74	62 ± 13.27	86 ± 4.06	62 ± 13.27
KNN	68.75 ± 4.84	65 ± 9.35	74.38 ± 3.64	65 ± 9.35

LR: Logistic Regression, RBF: Radial basis function, RF: Random Forest, KNN: K-Nearest Neighbour, SD: Standard deviation

3.2 FD Patients Without WMLs Versus Controls

Recursive feature elimination algorithm was performed on the 12 remaining features from filter method: cycle duration, cadence, loading, foot flat, peak swing, strike angle, minimum toe clearance, stride length variability, loading variability, lift-off angle variability, maximum heel clearance variability, and minimum toe clearance variability. Results are stated in Table 6 and Fig. 2. LR selected 8 features as the optimal number of gait characteristics with an F1 Score of 86.29%, SVM selected 6 features with an F1 Score of 74.29% and RF selected 10 features with an F1 Score of 68.88%. Table 6 presents the training and validation accuracies for the optimal models of each algorithm. LR presents the higher training and validation accuracies.

Table 6. Classification model accuracy for FD patients without WMLs versus controls

	Training accuracy % (Mean ± SD)	Validation accuracy % (Mean ± SD)
LR	96.44 ± 1.79	86.67 ± 19.44
SVM (Linear Kernel)	89.47 ± 6.79	78.33 ± 16.33
RF	95.53 ± 0.15	70 ± 2.67

Taking into account the contribution of each gait characteristic in the classification model (Fig. 2, right side), the common features were 5 (Top 5): loading, foot flat, minimum toe clearance, loading variability, and lift-off angle variability. The common gait characteristics from LR and SVM were 6 (Top 6): loading, foot flat, minimum toe clearance, stride length variability, loading variability, and lift-off angle variability. These Top 5 and Top 6, as well as the Top 3 from LR and SVM (stride length variability, loading variability, and lift-off angle variability), were evaluated with five classification models (LR, SVM Linear kernel, SVM RBF kernel, RF and KNN) to identify the optimal combination of gait

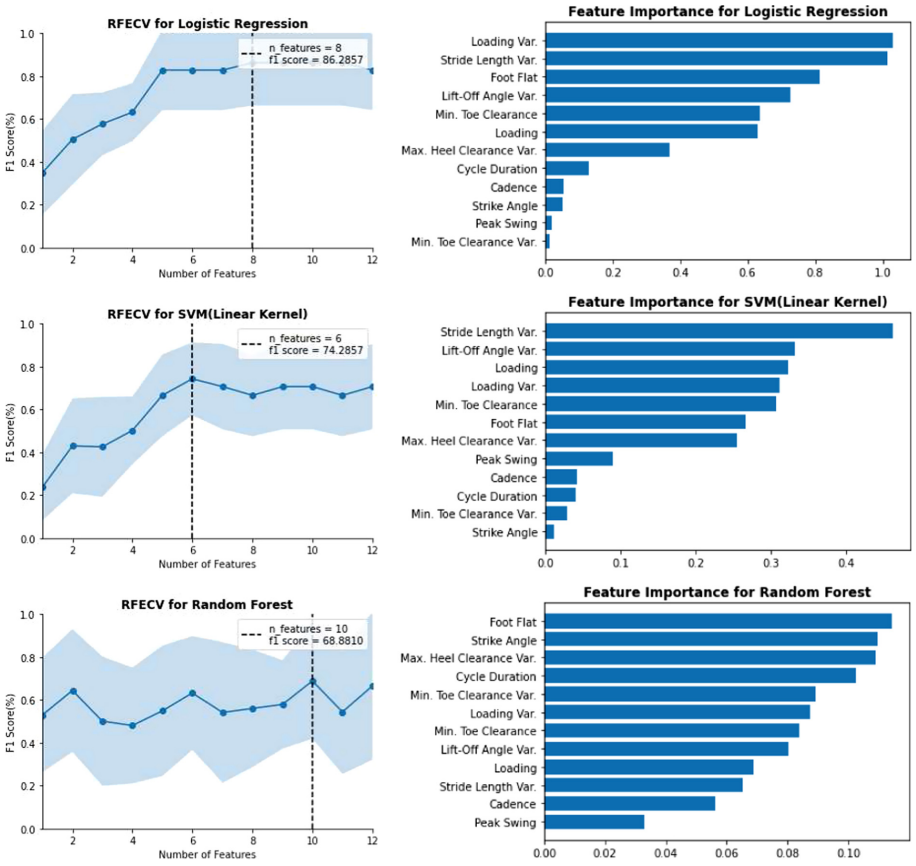

Fig. 2. F1 score versus number of features for selection of optimal numbers of gait characteristics (left) and feature importance results (right) obtained based on Logistic Regression, SVM Linear kernel and Random Forest in FD patients without WMLs vs. controls. Recursive features elimination was used through the 5-fold cross-validation (RFECV). Max.: maximum; Var.: variability; Min.: minimum.

characteristics and the classification model with better performance. Results are displayed in Table 7. With Top 5 SVM RBF kernel, RF and KNN achieved the highest validation accuracy of 76.67%. Looking at the Top 6, LR and SVM RBF kernel showed the highest validation accuracy of 83.33%. Using the Top 3 from LR and SVM, RF showed 78.33% validation accuracy, followed by the SVM Linear kernel, RF, and KNN with 75% validation accuracy, which KNN had the lower standard deviation. The validation accuracy of RF and KNN increased by reducing the feature set from 6 to 3, while the validation accuracy of LR, SVM Linear kernel, and SVM RBF kernel decrease. Overall, LR showed higher mean validation accuracy but also a higher standard deviation.

Table 7. Classification model accuracy for top common gait characteristics in FD patients without WMLs versus controls

	Top 5 common features (LR, SVM and RF)		Top 6 common features (LR and SVM)		Top 3 common features (LR and SVM)	
	Training accuracy (Mean ± SD)	Validation accuracy (Mean ± SD)	Training accuracy (Mean ± SD)	Validation accuracy (Mean ± SD)	Training accuracy (Mean ± SD)	Validation accuracy (Mean ± SD)
LR	85.76 ± 1.21	73.33 ± 17	97.35 ± 2.17	83.33 ± 18.26	79.47 ± 3.52	78.33 ± 12.47
SVM (Linear Kernel)	84.85 ± 1.92	70 ± 19.44	89.47 ± 6.79	78.33 ± 16.33	83.94 ± 4.56	75 ± 16.67
SVM (RBF kernel)	79.55 ± 4.07	76.67 ± 17	84.77 ± 4.75	83.33 ± 14.91	81.29 ± 5.09	70 ± 12.47
RF	89.32 ± 3.49	76.67 ± 13.33	91.91 ± 3.37	73.33 ± 17	90.08 ± 3.52	75 ± 16.67
KNN	82.27 ± 4.17	76.67 ± 13.33	83.11 ± 9.45	70 ± 16.33	84.85 ± 3.46	75 ± 7.45

LR: Logistic Regression, RBF: radial basis function, RF: Random Forest, KNN: K-Nearest Neighbour, SD: Standard deviation

4 Discussion

The present work aimed to identify discriminatory gait characteristics to distinguish FD with WMLs from controls and FD without WMLs from controls. Based on previous literature [1,7,17,20,21,26] different classification models were evaluated with different set of gait characteristics selected using a filter method follow by recursive feature elimination wrapper method with RF, SVM Linear kernel, and LR. Sixteen gait time series were obtained by two wearable sensors. All strides were normalized before developing any ML model according to previous studies [7,18,25]. Then, for each gait time series the mean and the standard deviation (as variability measure) were calculated obtaining 32 gait characteristics. From the feature selection analysis, foot flat, pushing, and maximum toe clearance 2 were identified as important characteristics to classify FD with WMLs. While stride length variability, loading variability, and lift-off angle variability, followed by loading, foot flat, and minimum toe clearance were identified as important gait characteristics to distinguish FD without WMLs from aged-matched healthy adults. Previous work [7] reveals that FD patients (with and without WMLs together) present lower percentages in foot flat and higher in pushing comparing with healthy adults.

For FD patients with WMLs versus controls, validation accuracy of 62–72% and a similar training accuracy of 69–85% was achieved thought the five selected classification models based on Top 3 gait characteristics, showing LR classifier the best performance with validation and training accuracy of 72% and 71%, respectively. With one more feature (foot flat variability) SVM with both Linear or RBF kernel also reveal good performance with an accuracy of 70% for validation and 74% for training. These results corroborated the hypothesis that the gait characteristics can be used to distinguish FD patients with WMLs from controls. This goes in line with the premise that gait is a final outcome of WMLs [22,23,27].

Surprisingly, in the FD patients without WMLs versus controls classification higher training and validation accuracies of 79–97% and 70–83% , respectively, were obtained based on Top 6, Top 5, or Top 3 selected features. LR and SVM RBF kernel classifier displayed the best performance based on Top 6, with an accuracy of 83% for validation and of 97% and 85% for training, respectively. By reducing the feature set for Top 3, overall validation accuracy of 70–78% was achieved, where for the RF and KNN classifiers the accuracy slightly increased to 75%. Similarly, in [21] an increase of the model accuracy was observed with feature reduction. Further, feature selection (reduction) plays an important role to deal with the problem of model overfitting, reduces training time, enhancing the overall ML performance and implementation. These results suggest that selected gait characteristics could be used as clinical features for supporting diagnoses of FD patients even without WMLs from younger ages since the mean age of these patients is 37.786 ± 10.48 years.

Due to the number of subjects involved in this study, all dataset was used in the training and validation of the models and any independent (external) dataset was used for checking the model performances. To test the robustness of classification models based on the selected gait characteristics further research with independent datasets is needed.

5 Conclusion

To the best of our knowledge, this is the first study that explores gait characteristics and their discriminate power in FD patients with WMLs and without WMLs from controls. For the discrimination of FD patients with WMLs the best model was built using LR with the variables foot flat, pushing, and maximum toe clearance 2, with an accuracy of 72%. In contrast, for the discrimination of FD patients without WMLs the best suited model was achieved using LR and SVM with RBF kernel based on the variables loading, foot flat, and minimum toe clearance, stride length variability, loading variability, and lift-off angle variability with an accuracy of 83%. The implications of WMLs on gait compromise in FD or predictive value of each kinematic gait variable remains still elusive, warranting further investigation with a more enriched cohort. Still, our findings are the first step to demonstrate the potential of machine learning techniques based on gait variables as a complementary tool to understand the role of WMLs in the gait impairment of FD. For future research, a larger sample size will be used to confirm and extend these findings.

References

1. Aich, S., Choi, K.W., Pradhan, P.M., Park, J., Kim, H.C.: A performance comparison based on machine learning approaches to distinguish Parkinson's disease from Alzheimer disease using spatiotemporal gait signals. Adv. Sci. Lett. **24**(3), 2058–2062 (2018)

2. Alcock, L., Galna, B., Perkins, R., Lord, S., Rochester, L.: Step length determines minimum toe clearance in older adults and people with Parkinson's disease. J. Biomech. **71**, 30–36 (2018)
3. Azevedo, O., et al.: Natural history of the late-onset phenotype of Fabry disease due to the p. F113L mutation. Mol. Genet. Metab. Rep. **22**, 100565 (2020)
4. Buechner, S., et al.: Central nervous system involvement in Anderson-Fabry disease: a clinical and MRI retrospective study. J. Neurol. Neurosurg. Psychiatry **79**(11), 1249–1254 (2008)
5. Burnham, K.P., Anderson, D.R.: Multimodel inference: understanding AIC and BIC in model selection. Sociol. Methods Res. **33**(2), 261–304 (2004)
6. Caramia, C., et al.: IMU-based classification of Parkinson's disease from gait: a sensitivity analysis on sensor location and feature selection. IEEE J. Biomed. Health Inform. **22**(6), 1765–1774 (2018)
7. Fernandes, C., et al.: Gait classification of patients with Fabry's disease based on normalized gait features obtained using multiple regression models. In: 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2288–2295. IEEE (2019)
8. Fernandes, C., et al.: Artificial neural networks classification of patients with Parkinsonism based on gait. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2024–2030. IEEE (2018)
9. Ferreira, F., et al.: Gait stride-to-stride variability and foot clearance pattern analysis in idiopathic Parkinson's disease and vascular parkinsonism. J. Biomech. **92**, 98–104 (2019)
10. Giugliani, R., et al.: A 15-year perspective of the Fabry outcome survey. J. Inborn Errors Metab. Screen. **4**, 1–12 (2016)
11. Gu, Q., Zhu, L., Cai, Z.: Evaluation measures of the classification performance of imbalanced data sets. In: Cai, Z., Li, Z., Kang, Z., Liu, Y. (eds.) ISICA 2009. CCIS, vol. 51, pp. 461–471. Springer, Heidelberg (2009). https://doi.org/10.1007/978-3-642-04962-0_53
12. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. J. Mach. Learn. Res. **3**(Mar), 1157–1182 (2003)
13. Katz, M.H.: Multivariable Analysis: A Practical Guide for Clinicians and Public Health Researchers. Cambridge University Press, Cambridge (2011)
14. Körver, S., Vergouwe, M., Hollak, C.E., van Schaik, I.N., Langeveld, M.: Development and clinical consequences of white matter lesions in Fabry disease: a systematic review. Mol. Genet. Metab. **125**(3), 205–216 (2018)
15. Kubota, K.J., Chen, J.A., Little, M.A.: Machine learning for large-scale wearable sensor data in Parkinson's disease: concepts, promises, pitfalls, and futures. Mov. Disord. **31**(9), 1314–1326 (2016)
16. Löhle, M., et al.: Clinical prodromes of neurodegeneration in Anderson-Fabry disease. Neurology **84**(14), 1454–1464 (2015)
17. Mannini, A., Trojaniello, D., Cereatti, A., Sabatini, A.M.: A machine learning framework for gait classification using inertial sensors: application to elderly, post-stroke and Huntington's disease patients. Sensors **16**(1), 134 (2016)
18. Mikos, V., et al.: Regression analysis of gait parameters and mobility measures in a healthy cohort for subject-specific normative values. PloS One **13**(6), 1–11 (2018)
19. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
20. Pradhan, C., et al.: Automated classification of neurological disorders of gait using spatio-temporal gait parameters. J. Electromyogr. Kinesiol. **25**(2), 413–422 (2015)

21. Rehman, R.Z.U., Del Din, S., Guan, Y., Yarnall, A.J., Shi, J.Q., Rochester, L.: Selecting clinically relevant gait characteristics for classification of early Parkinson's disease: a comprehensive machine learning approach. Sci. Rep. **9**(1), 1–12 (2019)
22. Snir, J.A., Bartha, R., Montero-Odasso, M.: White matter integrity is associated with gait impairment and falls in mild cognitive impairment. Results from the gait and brain study. NeuroImage: Clin. **24**, 101975 (2019)
23. Starr, J.M., et al.: Brain white matter lesions detected by magnetic resosnance imaging are associated with balance and gait speed. J. Neurol. Neurosurg. Psychiatry **74**(1), 94–98 (2003)
24. Thompson, C.G., Kim, R.S., Aloe, A.M., Becker, B.J.: Extracting the variance inflation factor and other multicollinearity diagnostics from typical regression results. Basic Appl. Soc. Psychol. **39**(2), 81–90 (2017)
25. Wahid, F., Begg, R., Lythgo, N., Hass, C.J., Halgamuge, S., Ackland, D.C.: A multiple regression approach to normalization of spatiotemporal gait features. J. Appl. Biomech. **32**(2), 128–139 (2016)
26. Wahid, F., Begg, R.K., Hass, C.J., Halgamuge, S., Ackland, D.C.: Classification of Parkinson's disease gait using spatial-temporal gait features. IEEE J. Biomed. Health Inform. **19**(6), 1794–1802 (2015)
27. Zheng, J.J., et al.: Brain white matter hyperintensities, executive dysfunction, instability, and falls in older people: a prospective cohort study. J. Gerontol. Ser. A: Biomed. Sci. Med. Sci. **67**(10), 1085–1091 (2012)