Improving the Lives of Multiple Sclerosis Patients through Wearable Technology

Neda Hassanpour, Adam Jenkins PhD, Charmaine Demanuele PhD

Introduction

Multiple Sclerosis (MS) is a chronic, inflammatory, demyelinating disease affecting the central nervous system. It disrupts information flow within the brain, and between brain and body. The underlying causes of MS are unknown, possibly involving gene-environment interactions. There is no cure for the disease, but treatments can help speed recovery from attacks, modify the course of the disease and manage symptoms.

Neuroimaging is often used for diagnosis, in follow-up visits, and clinical trails to assess the therapeutic effects of a new drug. Wearable devices may offer a complementary approach to neuroimaging: they can provide a continuous monitoring of the patient's function and quality of life, help detect new biomarkers of the disease, measure the long-term efficacy of new treatments, and empower the patient and healthcare providers.

Verily Dataset

As a joint study between Biogen, Verily, and Brigham and Women's Hospital (BWH), a dataset is collected to measure features of mobility, gait and dexterity with 9 wearable sensors in 25 ambulatory individuals with MS. The dataset includes 3 in-clinic visits and an 8-week at-home measurement of day-to-day activity between clinical visits 2 and 3. The dataset comprises MS Functional Component (MSFC) scores and patient demographics, as well as free-living features (such as sleep measures, and gait and motor metrics), and structured activity (such as psychomotor vigilance test, keyboard dexterity test, and mobility tasks).

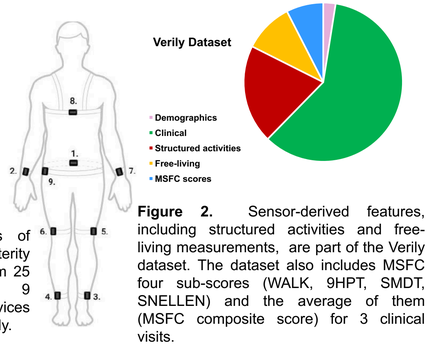


Figure 1. Features of mobility, gait and dexterity has been collected from 25 MS patients using 9 wearable sensor devices over eight weeks of study.

Figure 2. Sensor-derived features, including structured activities and free-living measurements, are part of the Verily dataset. The dataset also includes MSFC four sub-scores (WALK, 9HPT, SMDT, SNELLEN) and the average of them (MSFC composite score) for 3 clinical visits.

Project Definition and Research Objectives

This project is focused on the analysis of Verily dataset:

- **FeatureDay**: average value of the features for each day of study.
- **FeatureMedian**: features for the entire study period, in which the reported value for at-home features measured by sensors is the median of the observed day level values.

We seek to:

- Build an analysis pipeline that processes and visualizes the sensor data, and derives value from it.
- Asses the possibility of creating reliable models to predict MSFC scores for patients solely based on sensor device measurements.

June 2017 – August 2017

Features with High Correlation to MSFC Scores

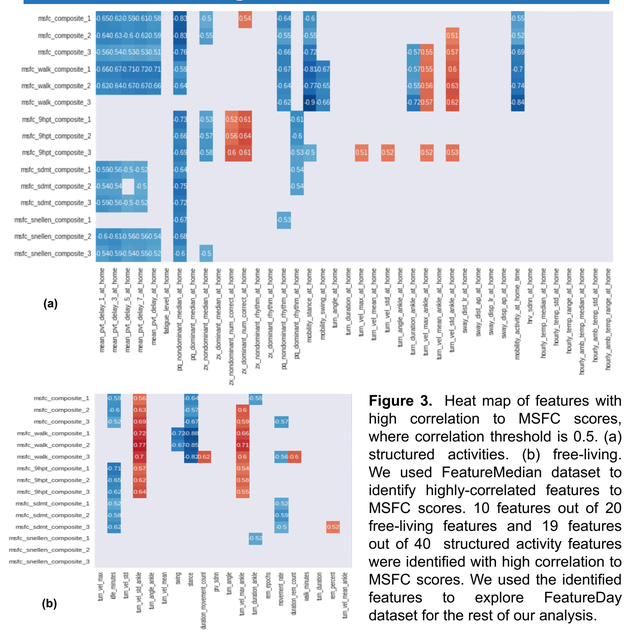


Figure 3. Heat map of features with high correlation to MSFC scores, where correlation threshold is 0.5. (a) structured activities. (b) free-living. We used FeatureMedian dataset to identify highly-correlated features to MSFC scores. 10 features out of 20 free-living features and 19 features out of 40 structured activity features were identified with high correlation to MSFC scores. We used the identified features to explore FeatureDay dataset for the rest of our analysis.

Rate of Change in Features and MSFC scores

Table 1. List of features, for which the rate of change is highly correlated with rate of change of MSFC scores in clinical visits 2 and 3. We estimated the rate of change for a feature by measuring the slope of a regression line that can fit best to the values of the feature throughout the course of the study. The rate of change of MSFC scores between clinical visits 2 and 3 is calculated using a similar technique. The thresholds used for correlation and p values are 0.5 and 0.05, respectively.

MSFC Score Component	Correlated Feature	Correlation Value	P value
msfc_composite	turn_duration_ankle	0.52	0.0188
msfc_walk	walk_duration	0.54	0.0170
msfc_9hpt	nondominant_rythm_pq_keyboad_test	0.57	0.0104
msfc_sdmt	--*	--	--
msfc_snellen	dominant_rythm_pq_keyboad_test	-0.77	0.0001

* no feature found

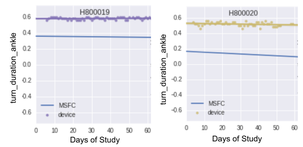


Figure 4. Visualizing the comparison of rate of change for the feature turn_duration_ankle and MSFC scores in two patients. The rate of change of this feature is identified highly correlated to rate of change in MSFC scores. The picture confirms the high correlation.

Variation in Features and MSFC Scores

MSFC Score Component	Correlated Feature	Correlation Value	P value
msfc_composite	idle_minutes	0.57	0.00793
	swing_duration_gait_cycle	0.56	0.00857
	stance_duration_gait_cycle	0.51	0.01991
	sleep_fraction_in_REM	0.64	0.00286
	swing_duration_walk	0.53	0.01447
	sleep_length	0.64	0.00305
	walk_duration	0.61	0.00371
msfc_walk	idle_minutes	0.58	0.00661
	swing_duration_gait_cycle	0.68	0.00082
	sleep_length	0.67	0.00149
	REM_sleep_length	0.62	0.00434
	sleep_fraction_in_REM	0.74	0.00024
	swing_duration_walk	0.72	0.00027
	walk_duration	0.76	0.00008
msfc_9hpt	--*	--	--
msfc_sdmt	idle_minutes	-0.54	0.01278
	turn_duration_ankle	-0.52	0.01824
msfc_snellen	dominant_rythm_pq_keyboad_test	0.72	0.00027

Table 2. List of features, for which the standard deviation of normalized feature values throughout the course of study is highly correlated to absolute change of MSFC scores in clinical visits 2 and 3. The thresholds used for correlation and p values are 0.5 and 0.05, respectively.

* no feature found

Stability and Variation of Features

Table 3. The most stable and variant features throughout the course of study. The features are identified by comparison of standard deviation of normalized feature values among all patients.

	Feature	Standard Deviation
Most stable (closest range)	dominant_rythm_pq_keyboad_test	0.0080
Most variant (widest range)	turn_max_angular_velocity_chest	0.1974

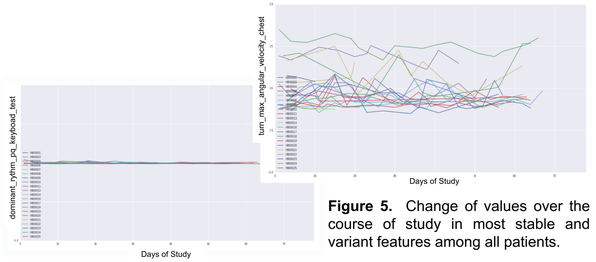


Figure 5. Change of values over the course of study in most stable and variant features among all patients.

Frequency of Change in Features and MSFC Scores

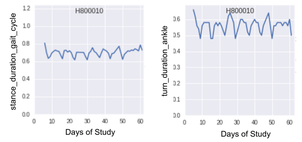


Figure 6. Some patients show obvious sine-wave pattern in change of feature values over the course of study. This picture shows an example of such sine-wave pattern in measurements for two specific features: turn_duration_ankle and stance_duration_gait_cycle.

Table 4. We extracted the most dominant frequency of change in feature values over the course of study using Fast Fourier Transform (FFT) and Power Spectral Density (PSD) analysis. We then looked for high correlations between the extracted frequencies and change in MSFC scores. Frequency of change of the feature sleep_length, is identified highly correlated to the change of MSFC composite score between clinical visits 2 and 3. The thresholds used for correlation and p values are 0.6 and 0.05, respectively.

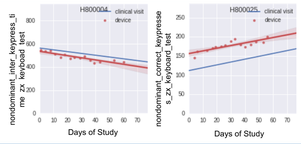
MSFC Score Component	Correlated Feature	Correlation Value	P value
msfc_composite	sleep_length	0.87	0.0210

At-home and Clinical Measurements

Table 5. List of structured activity features, for which the at-home measurements derived by sensors are capable of capturing the change that are measured in clinical visits. To identify these features, we calculated the correlation between the slope of regression lines that can fit best to the values of the features measured by sensors and clinical visits. The thresholds used for correlation and p values are 0.5 and 0.05, respectively.

Correlated Feature	Correlation Value	P value
turn_max_angular_velocity_ankle	0.56	0.0147
nondominant_inter_keypress_time_zx_keyboad_test	0.69	0.0012
nondominant_correct_keypresses_zx_keyboard_test	0.51	0.0228

Figure 7. Visualizing how at-home sensor measurements over the course of study are capable of capturing the change that is measured in clinical visits.



Multivariate Modeling for MSFC Score Prediction

Table 6. Summary of results on prediction of MSFC scores. We used normalized FeatureDay and FeatureMedian datasets as training data along with two targets: MSFC scores in clinical visit 3 (MSFC3), and the change in MSFC scores from clinical visit 2 to clinical visit 3 (MSFC3 - MSFC2). For the FeatureDay dataset, we considered each pair of a specific feature value and day of study as a new feature. We removed patients with more than 50% of sensor-derived data missing from the training datasets. We used Lasso regression model for predictions and 10-fold cross validation to evaluate the model by calculating Mean Squared Error (MSE).

Training Dataset	Target	#Samples	#Features	MSE	Regularization Coefficient in Lasso Model	Model Complexity (#Unique Features Selected for Model)
FeatureMedian	MSFC3	22	60	1.43	0.1162	2
FeatureDay	MSFC3	20	1826	0.67	0.0283	7
FeatureMedian	MSFC3 – MSFC2	22	60	0.20	0.0465	1
FeatureDay	MSFC3 – MSFC2	20	1826	0.17	0.0001	19

Multivariate Modeling for MSFC Score Prediction

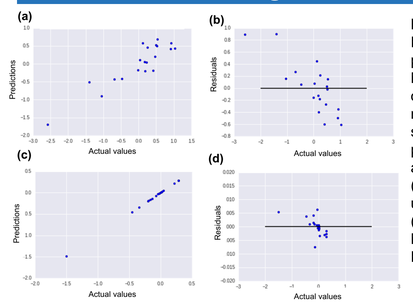


Figure 8. We found FeatureDay dataset promising in prediction of MSFC scores solely based on sensor device measurements. This picture shows comparison of predictions and residuals against actual values. (a) and (b) prediction of MSFC3 using FeatureDay dataset. (c) and (d) prediction of MSFC3 – MSFC2 using FeatureDay dataset.

Table 7. Summary of results on prediction of classified MSFC scores. We used FeatureDay and FeatureMedian datasets as training data along with 3-binned MSFC scores in clinical visit 3 (MSFC3_level) as the target. For the FeatureDay dataset, we considered each pair of a specific feature value and day of study as a new feature. We removed patients with more than 50% of sensor-derived data missing from the training datasets. We used Random Forest (RF) classifier for predictions and 10-fold cross validation to evaluate the model by calculating Mean Accuracy Score (MAS). We repeated this process on training datasets after applying Principle Component Analysis (PCA). We used grid search to find the best number of estimators in RF classifier and Principle Components (PCs) in PCA that lead to highest accuracy score.

Training Dataset	Target	#Samples	#Features	#Estimators	#PCs	MAS
FeatureMedian	MSFC3_level	22	60	10	NA*	0.81
FeatureDay	MSFC3_level	20	1826	5	NA	0.75
FeatureMedian_PCA	MSFC3_level	22	NA	10	2	0.88
FeatureDay_PCA	MSFC3_level	20	NA	5	17	0.86

* not applicable

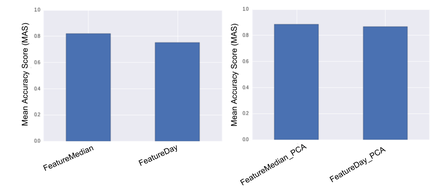


Figure 9. Comparing the mean accuracy score in prediction of classified MSFC score using different datasets. As expected, PCA transformed datasets result in better mean accuracy scores.

Discussion and Conclusion

This project is the first attempt to analyze the wearable segment of the Verily dataset. It will be highly beneficial to build an analysis pipeline that processes and visualizes this data, and derives value from it. We were able to identify certain features in the dataset that are promising for prediction of MSFC scores in order to evaluate the health state of the patients and effectiveness of drugs. In the long term, understanding the relationship between sensor metrics collected in real-world settings and clinical outcome measures is a step towards identifying novel, non-imaging biomarkers of the disease. This project could help inform treatment discovery as well as provide new markers for future clinical trials on neurodegenerative diseases.

Acknowledgements

A special thanks to my co-authors, Jason Tetrault, Effie Bruzik, Jane Rhodes, and all members of the Data Science team for all their help throughout this project.

Biogen