ANNALS of Clinical and Translational Neurology
Open Access
AMERICAN NEUROLOGICAL ASSOCIATION

RESEARCH ARTICLE

Biosensor vital sign detects multiple sclerosis progression

Kristen M. Krysko[1], Alireza Akhbardeh[2], Jennifer Arjona[1], Bardia Nourbakhsh[3], Emmanuelle Waubant[1], Pierre Antoine Gourraud[4,5] & Jennifer S. Graves[1,6]

[1]UCSF Weill Institute for Neurosciences, Department of Neurology, University of California San Francisco, San Francisco, CA
[2]Department of Radiology and Radiological Science, Johns Hopkins University, Baltimore, MD
[3]Department of Neurology, Johns Hopkins University, Baltimore, MD
[4]Nantes Université, CHU, INSERM, Centre de Recherche en Transplantation et Immunologie, UMR 1064, ATIP-Avenir, Nantes, France
[5]CHU de Nantes, INSERM, CIC 1413, Pôle Hospitalo-Universitaire 11: Santé Publique, Clinique des données, Nantes, France
[6]Department of Neurosciences, University of California San Diego, San Diego, CA

Correspondence
Jennifer S. Graves, UCSD ACTRI, 9452 Medical Center Drive, Box 0841, San Diego, CA 92037, USA. Tel: (858) 246-2384; Fax: 1-858-534-1078; Email: jgraves@health.ucsd.edu

Received: 13 July 2020; Revised: 21 August 2020; Accepted: 22 August 2020

Annals of Clinical and Translational Neurology **2021; 8(1): 4–14**

doi: 10.1002/acn3.51187

Abstract

Objective: To determine whether a small, wearable multisensor device can discriminate between progressive versus relapsing multiple sclerosis (MS) and capture limb progression over a short interval, using finger and foot tap data. **Methods**: Patients with MS were followed prospectively during routine clinic visits approximately every 6 months. At each visit, participants performed finger and foot taps wearing the MYO-band, which includes accelerometer, gyroscope, and surface electromyogram sensors. Metrics of within-patient limb progression were created by combining the change in signal waveform features over time. The resulting upper (UE) and lower (LE) extremity metrics' discrimination of progressive versus relapsing MS were evaluated with calculation of AUROC. Comparisons with Expanded Disability Status Scale (EDSS) scores were made with Pearson correlation. **Results**: Participants included 53 relapsing and 15 progressive MS (72% female, baseline mean age 48 years, median disease duration 11 years, median EDSS 2.5, median 10 months follow-up). The final summary metrics differentiated relapsing from secondary progressive MS with AUROC UE 0.93 and LE 0.96. The metrics were associated with baseline EDSS (UE $P = 0.0003$, LE $P = 0.0007$). While most had no change in EDSS during the short follow-up, several had evidence of progression by the multisensor metrics. **Interpretation**: Within a short follow-up interval, this novel multisensor algorithm distinguished progressive from relapsing MS and captured changes in limb function. Inexpensive, noninvasive and easy to use, this novel outcome is readily adaptable to clinical practice and trials as a MS vital sign. This approach also holds promise to monitor limb dysfunction in other neurological diseases.

Introduction

Accurate assessment of multiple sclerosis (MS) subtype and detection of disease progression remain clinical challenges. It is often only in several years retrospect that providers determine a patient has transitioned from relapsing to secondary progressive MS.[1] The distinction of MS subtype is important for clinical management and enrollment in trials.[2,3] There is also need for improved outcome measures that capture short-term progression to allow rapid testing of therapeutic agents in progressive MS trials.[4]

Leveraging biosensors from the gaming industry represents a novel approach to enhance the neurological exam. Sensor technology may offer the sensitivity and precision missing from current approaches to recording levels of neurological function. The most commonly studied sensor in MS involves use of accelerometers to measure step count as a marker of physical activity,[5] which correlates with Expanded Disability Status Scale (EDSS) in those with mild disability.[5–7] However, step count accuracy may be poor with altered gait,[8,9] and upper extremity function is not captured. Gyroscope sensors have been used to

assess balance,[5] whereas surface electromyogram (sEMG) sensors have not been as well-studied in MS.

The wearable multisensor MYO-band™ combines three sensors (accelerometer, gyroscope, sEMG). In a prior cross-sectional study, we demonstrated algorithms using this device had excellent reliability (intra-class correlation coefficients 0.80-0.87) and captured limb dysfunction with strong correlations with disability (Spearman correlation coefficient with EDSS 0.77 for UE and 0.82 for LE).[10] However, even more important than cross-sectional associations with gold standard outcomes is demonstrating a new outcome's ability to detect change over time. We aimed to develop a novel, agnostic (gold-standard independent) longitudinal algorithm to detect MS progression using the MYO-band sensor data. By not requiring a ground truth, this approach avoids the limitations of the EDSS. We specifically sought to objectively differentiate progressive from relapsing MS by capturing subtle progression over 1 year.

Methods

Study design

This is a longitudinal cohort study with prospective collection of demographic, clinical, and biosensor data during routine clinic visits approximately every 6 months to evaluate changes in limb function over time. The study period was from May 2016 to February 2018.

Study population

Adults 18 years and older with MS were recruited from the University of California San Francisco (UCSF) MS clinic with MS diagnosis based on 2010 McDonald criteria[11] confirmed by the study neurologist. All MS phenotypes were included, but individuals were excluded if limb motor strength was less than antigravity and the patient was unable to perform finger or foot taps. The Institutional Review Board at UCSF approved this study (15-18500). All participants provided written informed consent.

Procedures and measurements

Participants completed baseline questionnaires on study tablets to provide demographic and clinical information. An interval history form was completed at subsequent visits. Data were validated with medical records. Clinical diagnosis of MS subtype of relapsing remitting MS (RRMS), secondary progressive MS (SPMS) or primary progressive MS (PPMS)[12] was confirmed by the patient's neurologist and was used as the gold standard for MS subtype. For some analyses, SPMS and PPMS were combined into a progressive MS category, given clinical and pathological similarities.[13,14]

As a patient-centered outcome, participants completed the patient-reported telephone EDSS adapted for digital questionnaire use, which provides a score from 0 to 10 with higher scores indicating worse disability.[15] A Neurostatus-trained MS neurologist completed the physician-derived EDSS,[16] including functional system scores (FSS) and timed 25-foot walk (T25FW).

Participants wore the MYO-band during neurologic examination of finger and foot taps. For upper limb assessment, the small and expandable band was placed at the widest part of the upper forearm. They completed 20 index finger-thumb taps based on demonstration of 4 Hz and 4-inch amplitude taps. During seated foot tap testing, the device was placed at the widest part of the upper calf and they were instructed to tap the ball of the foot 20 times while leaving the heel of the foot on the ground based on demonstration of 4 Hz and 4-6-inch amplitude taps. The bands were placed in the same orientation on every participant. The tapping task was consistently demonstrated by the same examiner. The tapping test was chosen to allow mathematical modeling of repeated movement and has the advantage of using the same software and analytical approach for upper and lower limbs. Tapping tests have been shown to detect progression in MS.[17,18] The simplicity of tapping also allows the eventual generalizability of use by medical assistants and nonclinical personnel, as well as future application in other diseases.

Data acquisition

During assessments, the MYO was connected to an encrypted laptop via Bluetooth connection with immediate and postprocessing of signals. The data acquisition C++ code (Thalmic labs) was modified to export sensor data to text files in ASCII format, as described previously.[10]

Signal processing

Raw signals from the eight sEMG, three accelerometer, and three gyroscope channels were processed to reduce noise.[10] Each waveform was segmented to extract individual finger or foot tap events using a data-labeling algorithm in Matlab. Then, to mathematically characterize the sensor signals, seven waveform-based textural features were extracted for each tap for each of the three sensors to quantify limb performance. These included "mean value for each tap," "overall energy," "overall entropy," "haralick energy," "haralick correlation," "haralick contrast," and "haralick

homogeneity" (see Supplemental Methods).[19,20] In individual limbs of each participant, these seven textural features were each expressed in heatmaps for the three sensors separately, depicting the value of individual textural features during each tap for each limb. As an example, the Haralick energy textural feature for the right finger tap EMG data in a woman with PPMS and limb function worsening is shown in Figure 1. The change in the energy feature from blue to yellow represents more irregularity in the signal as the woman experienced more progression in that limb. Notably, the time it took to complete 20 taps also increased with the progression.

Derivation of scalar metrics of progression

Within-patient metrics of longitudinal change in limb function were created for each limb in each participant (see Supplemental Methods). To begin, in each participant and for each of the seven extracted textural features, the sum of absolute changes for each feature across the 20 tap events during each visit was calculated. Because this was calculated for each visit, this led to a matrix of absolute changes in textural features for each individual, with a matrix size that was the number of visits by the number of textural features plus the sum of the feature values. This type of matrix was created for each of the three sensors for each of the four limbs, leading to 12 matrices in each participant.

Next, to create a single scalar metric of longitudinal change in limb function, the change in all sensor signal waveform features across visits were combined. Signal waveforms were combined with a trapezoidal approach[21] in Matlab, which approximates integration over an interval by breaking the area into more easily computable areas. This takes into account the variable timing of real-world follow-up visits. These were then normalized by total follow-up time to take into account the variable duration of follow-up. This resulted in a limb-specific metric of change in limb performance over time within each participant. These metrics were created independently of the EDSS.

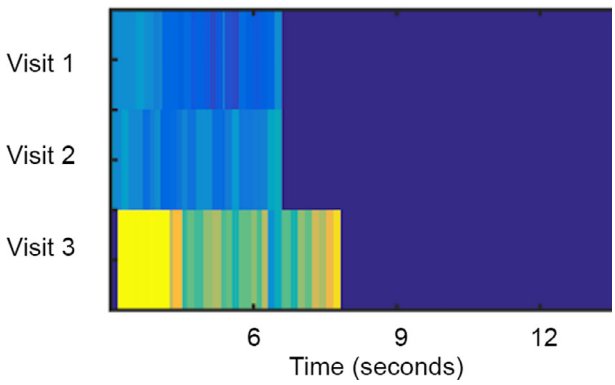

Figure 1. Haralick energy heatmap for finger tap EMG data for a patient with primary progressive MS. Figure legend: Heatmap for Haralick energy textural feature for the right finger tap EMG data, with three follow-up visits on the y-axis, with visits about 6 months apart over a 1-year period. This demonstrates the Haralick energy feature changes notably with change in heat map colors indicating more irregularity in tap movements over time in this individual with primary progressive MS, who developed inability to play piano with the right hand. EMG electromyogram; MS multiple sclerosis.

The limb-specific metrics of change for the right and left arm were summed to create a metric of UE change in function over time for each individual. Similarly, the limb-specific metrics of change for the right and left leg were summed to create a metric of LE change in function. The UE and LE metrics were combined to create an overall metric of change in limb function for each patient.

Statistical analysis

For descriptive analyses, mean and standard deviation, median and interquartile range or frequencies were reported as appropriate. Continuous variables were compared between MS subtypes with Kruskal–Wallis tests, and categorical variables were compared with chi-squared or Fisher exact tests.

Discrimination of MS subtype: Standard trapezoidal approach

UE and LE longitudinal metrics of within patient changes in limb function were compared between individuals with RRMS, PPMS, and SPMS using ANOVA. We also compared these metrics between progressive (SPMS and PPMS) and relapsing MS with t-tests, as well as logistic regression with calculation of area under the receiver operating characteristic (AUROC) curve. The metrics were also compared between SPMS and RRMS with logistic regression with calculation of AUROC. Multivariable logistic regression models were used to adjust for age, sex, and disease duration.

Discrimination of MS subtype: Machine learning generalized linear model (GLM)-based fusion approach

We next aimed to improve upon the MYO algorithm's AUROC performance in differentiating MS subtypes using the metrics calculated with the standard trapezoidal approach. This was done through use of generalized linear model (GLM)[22]-based fusion in which there was model

training using the gold standard of clinical MS subtype, as well as model cross-validation.

Data from each sensor in each limb were combined over time across visits with a trapezoidal approach normalized by follow-up time, as described above. However, before combining data from the different sensors, a GLM model was used to fuse input from all sensors in the right and left limbs, training the model with a label of MS subtype to improve performance. This was done individually for the UE, LE, and overall. The EDSS was not used in model training.

Given the small sample size, rather than splitting the data into test and validation sets and to avoid overfitting, we used the "leave-one-out" cross-validation approach.[23] Given n = 68, 68 separate times the model was trained on all except one participant, and a prediction was made for that participant, leading to a validation score for each individual for UE, LE, and overall from the model in which their data did not contribute. These scores were used to determine the AUROC for the UE, LE, and overall metrics. This approach was used to assess discrimination between progressive and RRMS, as well as between SPMS and RRMS. AUROCs based on this approach were compared to those derived with the standard trapezoidal approach.

Association with disability

We compared the digital within patient UE and LE metrics of change to baseline and changes in physician and patient-reported EDSS, pyramidal and cerebellar FSS, and T25FW with Pearson correlation as distributions were approximately normal. For exploratory purposes, we also evaluated the performance of baseline EDSS or change in EDSS alone in their ability to discriminate MS subtype. We explored characteristics of patients with no change in EDSS, but digital signature metrics more than 1 standard deviation (SD) higher than the mean in either the UE or LE for exploratory purposes.

All tests were two-sided with an alpha of 0.05. Analyses were performed using STATA 15 (College Station, TX) and Matlab. Model diagnostics showed no concerning departure from model assumptions and model fit was adequate.

Results

The cohort included 68 patients with mean baseline age of 48 years (72% female). Other characteristics are described in Table 1. As expected, those with progressive MS were older than RRMS, and had higher baseline EDSS. Those with SPMS had longer disease duration than other MS subtypes. These differences were accounted for in the multivariable models below. Median follow-up was 10 months (interquartile range (IQR) 6 to 15) with a median of three visits each (IQR 2 to 4).

Derivation of scalar metrics of progression and discrimination of MS subtype

As described in the Methods, calculating change over time in textural features of sensor waveform data captured changes in the function of each limb. The final unitless scalar metric of limb progression ranged from 328 to 1799 in the LE (n = 64) and 590 to 2305 in the UE (n = 66), with higher numbers indicative of more progression. The UE and LE longitudinal metrics differed by MS subtype (LE: RR 823, PP 986, SP 1440, $P < 0.001$; UE: RR 1065, PP 1221, SP 1658, $P < 0.001$) (Fig. 2). The scalar metrics also differentiated RR from both types of progressive MS (LE: RR 823, progressive 1126, $P = 0.008$; UE: RR 1065, progressive 1396, $P = 0.012$). The higher magnitude values in the progressive participants indicate a greater change across textural features and sensors over time.

Standard trapezoidal analytical models

Using a trapezoidal function in Matlab that accounted for different follow-up times for different patients, an AUROC analysis, and a standard logistic regression model, we evaluated the performance of the UE, LE, and overall longitudinal sensor metrics' abilities to distinguish between relapsing and progressive MS (SPMS and PPMS). The metrics discriminated RRMS from the combined progressive MS group, with AUROC 0.80 (95% CI 0.62 to 0.98) for LE, 0.67 (95% CI 0.47 to 0.87) for UE, and 0.84 (95% CI 0.68 to 0.99) for the overall 4-limb metric (Fig. 3). In an unadjusted regression analysis, for every 1 standard deviation (SD) unit increase in LE metric, there was 2.7 times the odds (95% CI 1.4-5.3, $P = 0.003$) of progressive MS. In unadjusted analysis, for every 1 SD unit increase in the UE metric, there was 2.8 times the odds (95% CI 1.4-5.3, $P = 0.002$) of progressive MS. The magnitude of these associations remained similar after adjustment for age, disease duration and sex (Table 2).

Most relevant to clinical practice, the metrics performed even better in differentiating SP from RR MS, with an AUROC 0.96 (95% CI 0.90 to 1.00) for LE, 0.93 (95% CI 0.86 to 1.00) for UE, and 0.96 (95% CI 0.90 to 1.00) for the overall metric (Fig. 3). In unadjusted analysis, for every 1 SD unit increase in LE metric, there was 6.7 times the odds (95% CI 1.7-26.7, $P = 0.007$) of SPMS than RRMS and for every 1 SD unit increase in UE metric, there was 7.0 times the odds (95% CI 1.9-24.9, $P = 0.003$) of SPMS. These associations remained robust

Table 1. Baseline characteristics of participants (n = 68)

Characteristic	All (n = 68)	Relapsing remitting MS (n = 53)	Secondary progressive MS (n = 6)	Primary progressive MS (n = 9)	*P*-value
Age, mean years (SD)	48.3 (12.1)	45.7 (11.8)	57.3 (11.8)	57.4 (5.9)	0.0018[a]
Female sex, n (%)	49 (72.1%)	43 (81.1%)	2 (33.3%)	4 (44.4%)	0.007[b]
Disease duration, median years (range)	10.5 (0.1-44.0)	10.0 (0.1-44.0)	25.5 (10.0-40.0)	7.0 (3.0-14.0)	0.027[a]
DMT, n (%)					0.009[c]
None	19 (27.9%)	10 (18.9%)	3 (50%)	6 (66.7%)	
Interferon-beta-1a	4 (5.9%)	3 (5.7%)	1 (16.7%)	0	
Glatiramer acetate	12 (17.6%)	12 (22.6%)	0	0	
Dimethyl fumarate	12 (17.6%)	12 (22.6%)	0	0	
Fingolimod	5 (7.4%)	5 (9.4%)	0	0	
Natalizumab	10 (14.7%)	9 (17.0%)	0	1 (11.1%)	
Rituximab	6 (8.8%)	2 (3.8%)	2 (33.3%)	2 (22.2%)	
EDSS, median (range)	2.5 (0.0-7.0)	2.0 (0.0-7.0)	6.0 (3.0-7.0)	4.0 (3.0-6.5)	0.0001[a]
Self-reported EDSS, median (range)	2.0 (0.0-7.0)	2.0 (0.0-7.0)	5.5 (2.0-6.5)	4.5 (2.0-6.5)	0.0002[a]
25-foot walk time, median seconds (range)	4.1 (2.8-30.5)	3.9 (2.8-30.5)	7.8 (5.2-16.5)	5.2 (4.0-19.6)	0.0004[a]

MS multiple sclerosis; SD standard deviation; DMT disease-modifying therapy; EDSS Expanded Disability Status Scale.
Continuous variables were compared between MS subtypes with the Kruskal–Wallis test,[a] and categorical variables compared with the chi-squared test[b] or Fisher exact test[c].

with adjustment for age, disease duration and sex (Table 2).

Machine learning GLM-based fusion analytical models

With "leave-one-out" cross-validation of the GLM-based fusion approach, the AUROC for discrimination between MS subtypes improved over the standard trapezoidal approach, with statistically significant improvement for the UE metric in differentiating progressive from relapsing MS, which had the poorest AUROC performance with the standard trapezoidal approach. With the GLM-based fusion approach, for discrimination of the combined progressive group from relapsing MS, AUROC was 0.93 (95% CI 0.83 to 1.00) for UE, 0.84 (95% CI 0.69 to 0.99) for LE, and 0.85 (95% CI 0.71 to 1.00) for the overall metric. For discrimination of SPMS from RRMS, AUROC was 0.92 (95% CI 0.75 to 1.00) for UE, 0.99 (95% CI 0.95 to 1.00) for LE, and 0.99 (95% CI 0.95 to 1.00) for the overall metric (Table 2).

Association with disability

Although the analytical approach was not dependent on a change in EDSS, we anticipated that those with higher baseline EDSS would be at greater risk for progression over time. Thus, we hypothesized that higher baseline disability would be associated with greater magnitudes of the new biosensor metrics of progression. Higher values of both the UE and LE sensor-derived metrics were associated with baseline disability as measured by the physician and self-reported EDSS, pyramidal, and cerebellar FSS and T25FW (Fig. 4, Table 3).

As expected, there was little change in EDSS over the median follow-up of 10 months (only 29% had any change in EDSS). The median change in EDSS was 0.0 for RRMS (range = −2.0 to + 1.5), 0.0 (0.0 to +4.0) in SPMS, and 0.0 (−1.0 to 0.0) in PPMS. These few changes in EDSS were not due to relapses, as only 1 patient with RRMS had a relapse during the follow-up period, and that individual had stable EDSS. With little signal change for EDSS over the short time interval, it was not expected to see strong longitudinal associations with the new biosensor metrics. Nonetheless, there was a statistically significant correlation between the LE metric and change in EDSS (Fig. 4, Table 3).

In this cohort, the baseline gold standard EDSS discriminated between RRMS and progressive MS (AUROC 0.88, 95% CI 0.80 to 0.96) and between RRMS and SPMS (AUROC 0.90, 95% CI 0.77 to 1.00). On the other hand, as expected, change in EDSS performed poorly in discriminating between MS subtypes (RRMS versus progressive MS: AUROC 0.54, 95% CI 0.40 to 0.68; RRMS versus SPMS: AUROC 0.61, 95% CI 0.35 to 0.87).

We explored in more detail those individuals with stable EDSS, but the UE or LE metric of limb progression value was ≥1 SD higher than the mean value. Of the eight individuals with stable EDSS but the UE progression metric ≥1 SD than the mean, the majority had progressive MS (3 SPMS, 2 PPMS), and the three with RRMS had

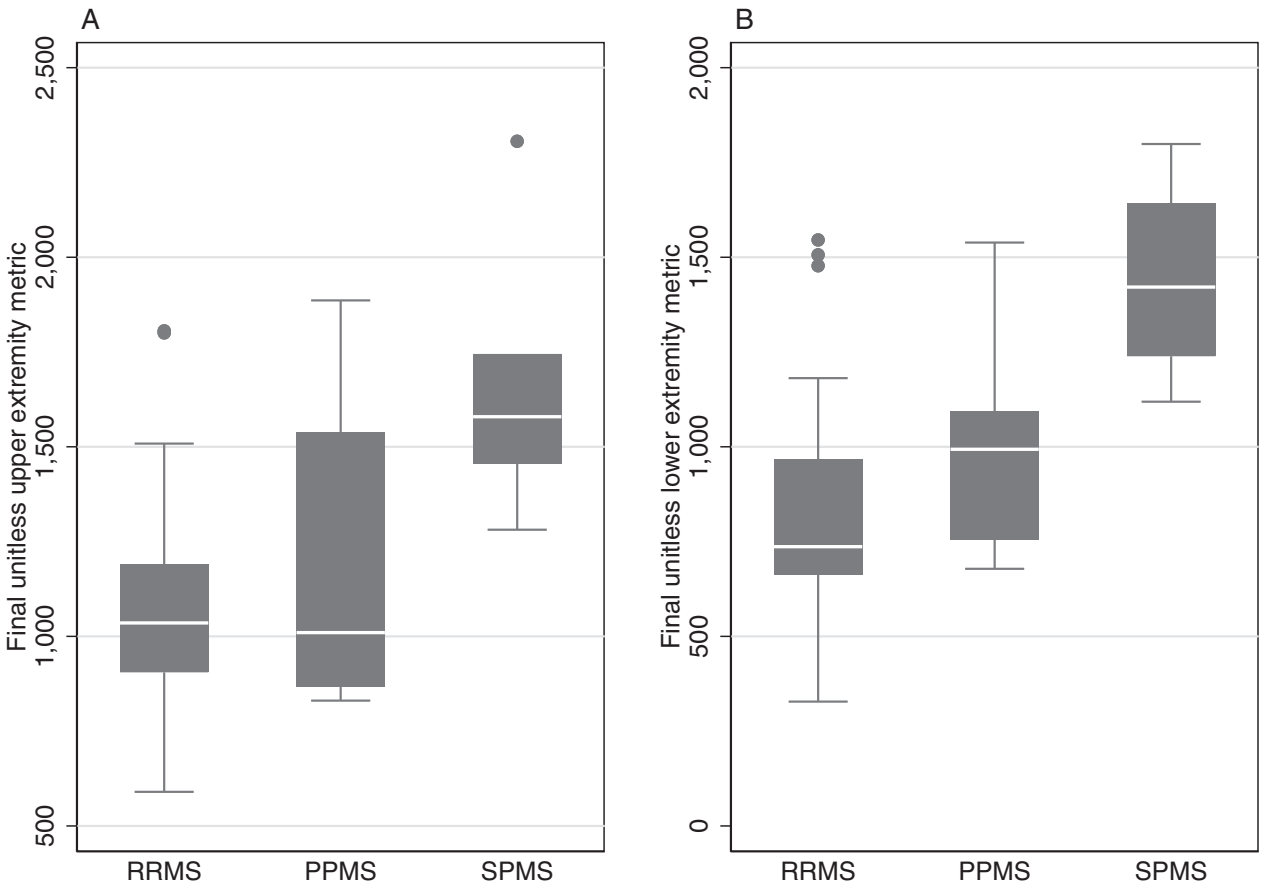

Figure 2. Final scalar metric for the upper extremity (A) and lower extremity (B) by MS subtype. Figure legend: The final scalar metric differentiated multiple sclerosis subtype in the upper extremity ($P < 0.001$) and lower extremity ($P < 0.001$), with higher scores in those with progressive than relapsing MS. MS multiple sclerosis; PPMS primary progressive MS; RRMS relapsing remitting MS; SPMS secondary progressive MS.

baseline EDSS scores of 2.0, 3.0, and 6.0. Of the four individuals with stable EDSS but the LE progression metric ≥1 SD than the mean, the majority had progressive MS (2 SPMS, 1 PPMS), and the one with RRMS had an EDSS of 6.0. We suspect the metric suggests subclinical progression not detected by EDSS in these individuals.

Discussion

We have developed a novel longitudinal algorithm to capture MS progression in individual limbs over the short time course of a year. With potential applications beyond MS, the multisensor hardware is small, wearable and inexpensive, and the algorithm is independent of a disease specific ground truth. Our results suggest we can detect limb progression not captured by gold standard MS disability metrics. The testing paradigm is brief, only taking a few minutes to complete, and fits within the general clinical flow of a neurology visit. Our algorithm could enhance the physical exam as a novel neurological vital sign, and be applied to data collected during routine clinic and research visits. Further development of commercial grade software will allow real-time generation of UE, LE, and overall limb progression scores for the clinician and patient. Analytics could be performed with immediate output of a vital sign value for each limb and no need for signal processing on-site. These data could be used to classify MS subtype objectively and efficiently for more appropriate clinical management and clinical trial assignments. This tool could also provide enhanced adjudication of treatment effects for therapeutics in development for progressive MS. Our approach could also be applied to stroke, Parkinson's disease, motor neuron disease, and other neurological conditions to monitor limb function over time.

Conventionally, diagnosis of MS subtype is largely retrospective and based on clinician impression. The distinction between RRMS and SPMS during the transition is particularly difficult.[1] EDSS-based algorithms can allow

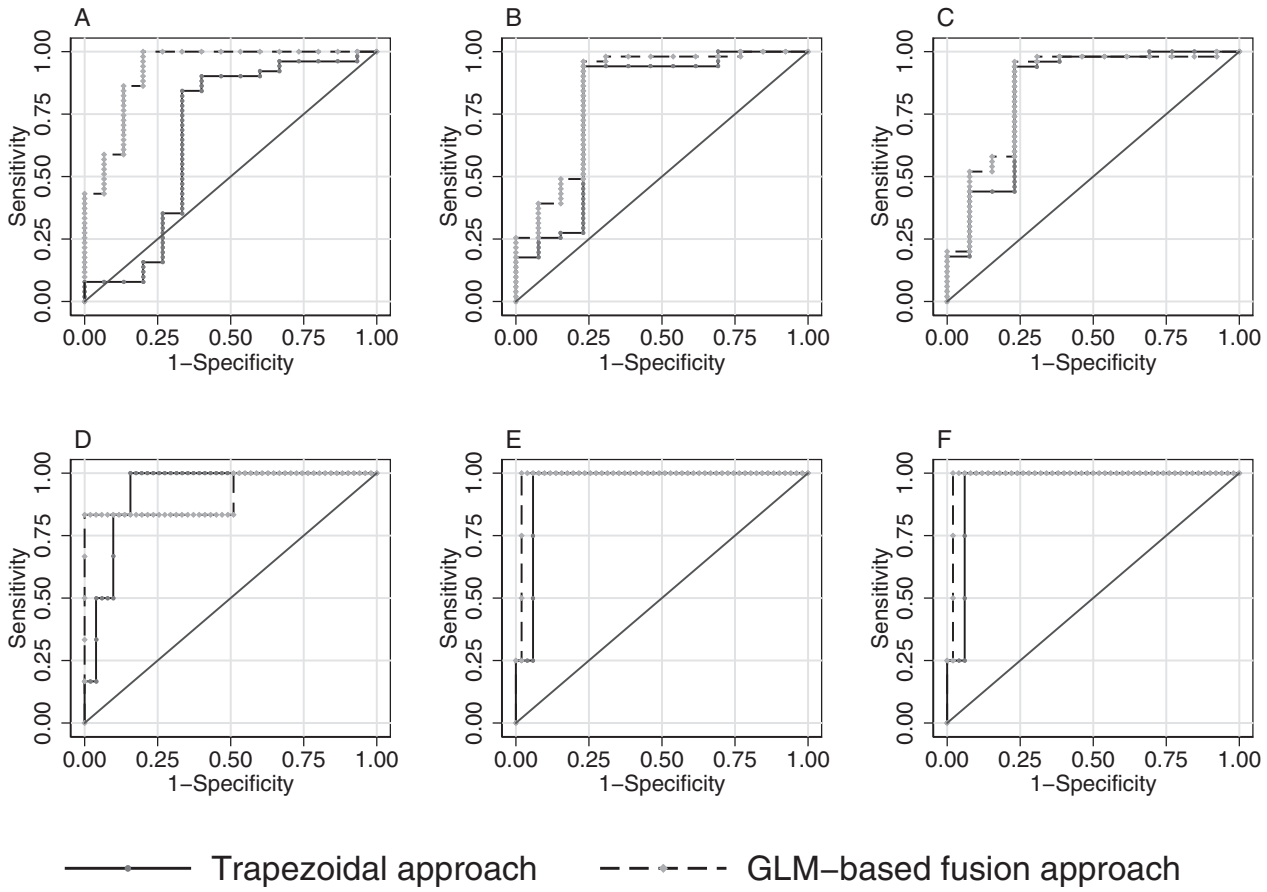

Figure 3. ROC curves discriminating MS subtype with upper and lower extremity metrics, and the overall metric. Figure legend: Discrimination of all progressive types versus relapsing MS with the upper extremity metric (A), lower extremity metric (B), and overall combined upper and lower extremity metric (C). Discrimination of secondary progressive from relapsing MS with the upper extremity metric (D), lower extremity metric (E) and overall combined upper and lower extremity metric (F). These were calculated with the standard trapezoidal approach and the GLM-based fusion approach. GLM generalized linear model; MS multiple sclerosis; ROC receiver operating characteristic.

earlier diagnosis of SPMS.[24] However, these rely on the EDSS, which is problematic given lack of sensitivity to short-term change in disability, inter-rater variability, lengthy assessments, and reliance on ambulatory dysfunction at higher scores.[25,26] Our study was similar to prior work in that baseline EDSS was higher in progressive MS than RRMS, but change in EDSS was unable to discriminate between MS subtypes. In contrast, the longitudinal MYO progression scores were able to distinguish among MS subtypes.

Wearable biosensors hold excellent promise to monitor disability in MS. While laboratory biomarkers have also been proposed to distinguish MS subtype and monitor progression,[27,28] digital biomarkers have the advantages of immediate feedback on disease state, ability to employ them directly in clinics, and the potential to distribute them into patient's homes for remote monitoring. Previously, biosensing studies in MS have primarily included accelerometers to capture walking disability.[5,29] The approach of using sensors to enhance the exam requires less time commitment from the patient and provider. The use of the multisensor MYO device is also novel as it captures upper and lower extremity disability with the same device and software and provides a disability estimate for each limb individually.[10] This is in contrast and complementary to other biosensors that estimate physical activity or gait function, as we are able to provide a performance metric for each limb over time. Beyond finger and foot taps, the MYO device could also be extended to evaluate other movements assessed during the neurological examination.

Those with greater disability had higher values of our longitudinal metric, which is expected as those with higher EDSS are expected to be more prone to insidious progression. There was also an association between

Table 2. Distinguishing progressive from relapsing multiple sclerosis with the final scalar metrics using the trapezoidal approach and GLM-based fusion approach

	Unadjusted			Adjusted[b]			Trapezoidal approach	GLM fusion approach	AUROC comparison
	OR[a]	95% CI	*P*-value	OR[a]	95% CI	*P*-value	AUROC (95% CI)	AUROC (95% CI)	*P*-value[c]
All Progressive vs. Relapsing MS									
All 4 limb metric	-	-	-	-	-	-	0.84 (0.68 to 0.99)	0.85 (0.71 to 1.00)	0.67
Upper extremity	2.8	1.4 to 5.3	0.002	3.0	1.3 to 6.9	0.011	0.67 (0.47 to 0.87)	0.93 (0.83 to 1.00)	0.022
Lower extremity	2.7	1.4 to 5.3	0.003	2.3	0.99 to 5.5	0.053	0.80 (0.62 to 0.98)	0.84 (0.69 to 0.99)	0.084
Secondary progressive vs. Relapsing MS									
All 4 limb metric	-	-	-	-	-	-	0.96 (0.90 to 1.00)	0.99 (0.95 to 1.00)	0.20
Upper extremity	7.0	1.9 to 24.9	0.003	15.5	1.4 to 170.6	0.025	0.93 (0.86 to 1.00)	0.92 (0.75 to 1.00)	0.88
Lower extremity	6.7	1.7 to 26.7	0.007	5.7	0.89 to 37.0	0.066	0.96 (0.90 to 1.00)	0.99 (0.95 to 1.00)	0.20

OR odds ratio; CI confidence interval; AUROC area under the receiver operating characteristic curve; GLM generalized linear model.

[a]Per 1 standard deviation unit increase in the final scalar metric.

[b]Adjusted for baseline age, sex, and disease duration.

[c]These p-values compare AUROC methods, comparing whether there is a difference between the trapezoidal and GLM fusion approach. The GLM approach improved UE metric performance for the discrimination of progressive and relapsing MS.

change in EDSS and the LE but not UE biosensor metric. This is expected as EDSS is strongly dependent on ambulatory function, and thus change in EDSS is more reflective of lower limb function. The EDSS is much less likely to capture change in upper limb function at higher scores. On the other hand, the magnitudes of the new biosensor UE metric appeared to capture worsening of arm and hand function over the course of the study. Given the limitations of the EDSS, our individual limb-based metrics may be more sensitive to progression. There were several participants with no change in the EDSS who had high values of the longitudinal metric of change, mainly with progressive MS, suggesting subtle progression not captured by EDSS.

Limitations include lack of a perfect or pathology-based gold standard for MS subtype or change in disability over short intervals to compare with our new metrics. Many patients did not demonstrate progression on the EDSS or T25FW, making it difficult to prove that changes we detected with the MYO were representative of true pathophysiological progression or predictive of clinical worsening. This limitation is not unique to our study, as it would occur in the evaluation of any new metric developed to improve upon an imperfect gold standard such as detection of progression by EDSS. Given these limitations in measuring disability progression over a short interval, we focused evaluation in this pilot study on the ability to distinguish the progressive from relapsing MS phenotype to ensure that the magnitude of the agnostic longitudinal metrics was meaningful. Given the proof of concept nature of this study with emphasis on signal processing algorithm development and limited time during routine clinic visits, we did not include other UE-specific disability measures such as the 9-hole peg test or force plate motor assessments, or other patient-reported outcomes to compare with the MYO metrics. These are of interest in future studies, which will focus on evaluating the ability of these metrics to capture clinically meaningful disability progression in a larger sample over a longer time period. This pilot study also did not have the scope to include MRI lesion load or atrophy measures, optical coherence tomography metrics, or serum neurofilament light chain, which could be considered in future larger studies for biomarker comparisons. Additionally, the imperfect gold standard of MS subtype based on clinical assessment could impact the metrics' performance in classifying MS subtype. Despite this, we were able to demonstrate associations between the novel metrics, disability measures and MS subtype. Additionally, the sample size was relatively small with a limited number with progressive MS in this pilot study, which is a limitation particularly given our primary aim involved differentiating progressive from relapsing MS. Despite the small sample size, we were able to demonstrate statistically significant differences in the longitudinal metrics between MS subtypes. Finally, although this was a single-center study, a broad range of patients were included with all MS subtypes represented and a wide range of disability and disease duration.

Strengths of our study include the use of the MYOband, which is inexpensive, comfortable, and quick and easy to use during a standard neurologic exam with minimal additional time required. Ongoing monitoring during daily activities is not required. Additionally, this single device and algorithm are used to evaluate both arm and leg function, and nonclinical personnel could perform the

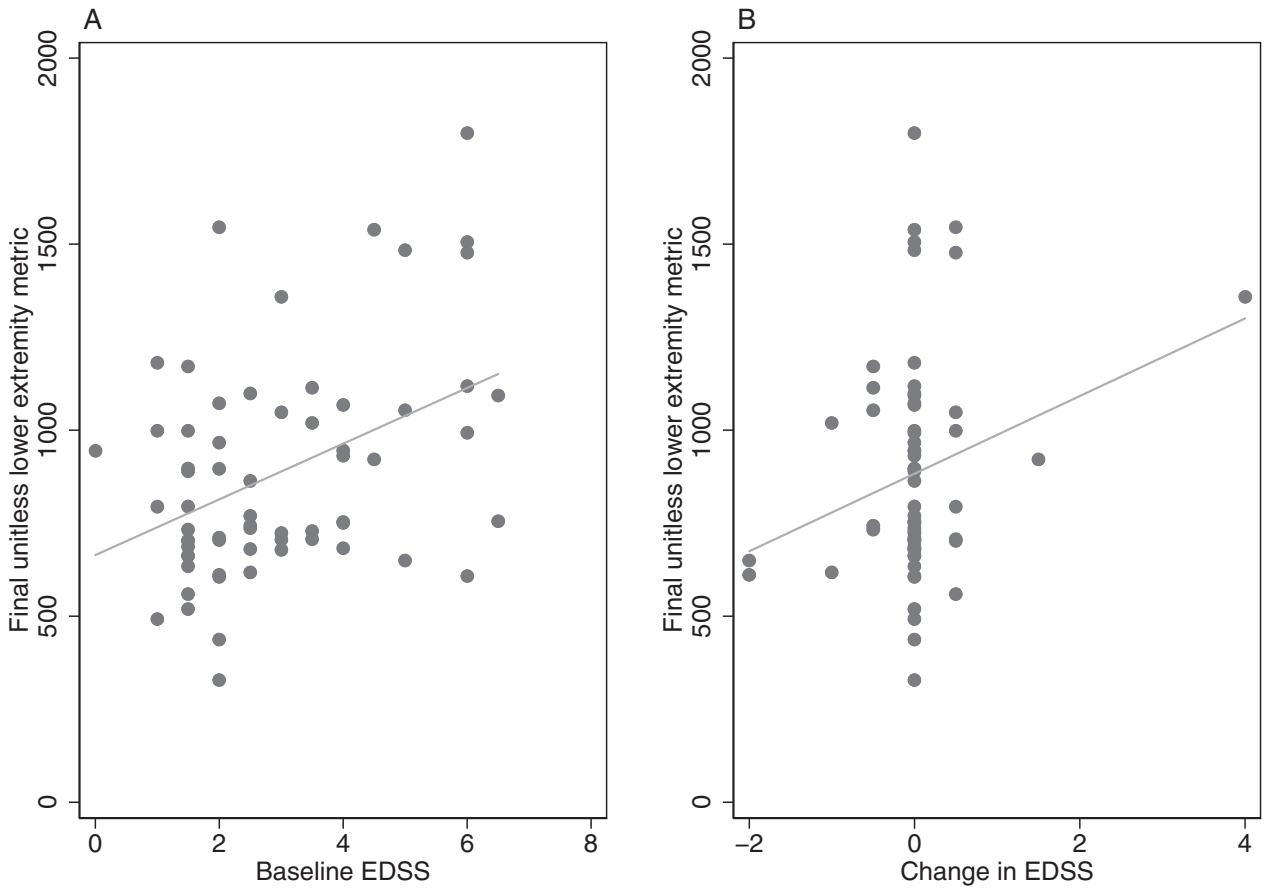

Figure 4. Association of the lower extremity metric with baseline (A) and change (B) in disability. Figure legend: The lower extremity metric was associated with baseline Expanded Disability Status Scale (EDSS) (r = 0.41, $P = 0.0007$) and with change in EDSS (r = 0.25, $P = 0.048$) although several individuals had no change in EDSS, but high values of the metric suggesting progression not detected by the EDSS. Positive values for change in EDSS indicate worsening disability.

assessment. Furthermore, our algorithm adjusts for variability in follow-up time, which is important for use in real-world practice. The algorithm was optimized with a machine learning approach. The best discrimination was between RR and SPMS, which is clinically the most relevant comparison given individual patients transition between these MS subtypes, and it is difficult to determine when the transition occurs. This distinction will be more important as therapies become available for SPMS.[3,30] It is important to note that this device does not differentiate between inflammatory disease activity and disability worsening due to progression, and as such may not guide clinical decisions regarding the use of currently available disease-modifying therapies that primarily target inflammatory activity.

We have developed an algorithm using a small, nonobtrusive wearable multisensor device that can detect change in upper and lower limb function over a short interval. This algorithm can be used to classify MS subtype and detect subtle progression not detected by standard disability assessments, which could be used to monitor treatment effectiveness. While in this proof of concept study we performed signal-processing on-site, for widespread dissemination for clinical practice or trials, commercial grade software would be utilized for real-time calculation of limb progression scores. This would allow immediate output of these MS vital signs, which would be available for use during clinical evaluation, similar to blood pressure and heart rate as measures of cardiovascular function. Future studies are required to validate the algorithm in an independent, larger sample collected at multiple sites including more individuals with progressive MS. This is a promising tool to monitor progression in clinical trials. Furthermore, this tool is not limited to use in MS and could be applied to a broader range of neurologic diseases affecting limb function, including stroke, Parkinson's disease, and motor neuron disease. Future study will involve evaluation of this algorithm for

Table 3. Association between the final upper and lower extremity MYO longitudinal metrics and baseline and change in disability

	Upper extremity metric		Lower extremity metric	
	Pearson correlation coefficient	*P*-value	Pearson correlation coefficient	*P*-value
Baseline EDSS	0.43	0.0003	0.41	0.0007
Baseline pyramidal FSS	0.39	0.0014	0.33	0.0071
Baseline cerebellar FSS	0.40	0.0008	0.45	0.0002
Baseline self-reported EDSS	0.40	0.0009	0.38	0.0021
Baseline 25-foot walk time	0.31	0.022	0.38	0.0046
Change in EDSS	0.12	0.33	0.25	0.048
Change in pyramidal FSS	0.12	0.34	0.19	0.14
Change in cerebellar FSS	0.04	0.73	0.03	0.79
Change in self-reported EDSS	0.03	0.84	0.01	0.91
Change in 25-foot walk time	-0.06	0.64	-0.02	0.86

EDSS Expanded Disability Status Scale; FSS functional system score.
Few individuals had change in standard disability metrics over the course of the study (only 20/68 had change in EDSS, 10/68 change in pyramidal FSS, 11/68 change in cerebellar FSS, and 36/68 change in self-reported EDSS; median change in T25FW was +0.11 seconds (IQR −0.26 to +0.57)).

monitoring limb dysfunction in these other neurological diseases.

Author Contributions

KMK had a role in study design; designed the study analysis plan; major role in acquisition of data; analyzed and interpreted the data; drafted the manuscript for intellectual content. AA designed the signal processing algorithm; analyzed and interpreted the data; revised the manuscript for intellectual content. JA had a role in study design; major role in acquisition of data; revised the manuscript for intellectual content. BN had a role in study design; role in acquisition of data; revised the manuscript for intellectual content. EW had a role in study design; role in acquisition of data; revised the manuscript for intellectual content. PA had a role in study conceptualization; revised the manuscript for intellectual content. JSG designed and conceptualized the study; obtained funding; designed the signal processing and statistical analysis; major role in acquisition of data; interpreted the data; drafted and revised the manuscript for intellectual content.

Acknowledgments

To the dedicated patients who participated in this study. We thank Dr. Douglas Goodin, Thomas Capenito, and Adam Santaniello for their assistance with the patient-reported EDSS questionnaire and digital adaptation. Study funding included the UCSF CTSI pilot grants program and an Investigator initiated grant from Genentech. Dr. Kristen Krysko is funded by a Sylvia Lawry Physician Fellowship through the National Multiple Sclerosis Society (FP-1605-08753 (Krysko)).

Conflicts of Interest

Dr. Kristen Krysko is funded by a Sylvia Lawry Physician Fellowship through the National Multiple Sclerosis Society (FP-1605-08753 (Krysko)). She also has fellowship funding through Biogen. Dr. Alireza Akhbardeh has no disclosures. Ms. Jennifer Arjona has no disclosures. Dr. Bardia Nourbakhsh reports personal fees from Jazz Pharmaceutical and grants from Genentech, outside the submitted work. Dr. Emmanuelle Waubant reports personal fees from DBV, Jazz Pharmaceuticals, Emerald, outside the submitted work. Dr. Pierre Antoine Gourraud received consulting fees or sponsored research from major pharmaceuticals companies all dealt with through academic pipelines: Merieux, Biogen, Merck, Methodomics, WeData, Boston Scientific, AstraZeneca, Cook. He was also founder (2008) www.methodomics.com and co-Founder (2018) www.wedata.science. Dr. Jennifer Graves received grants from Genentech during the conduct of the study. She received grants from Biogen, personal fees from Novartis, and grants from Octave outside the submitted work.

Data availability statement

Software and de-identified data may be available upon reasonable request to the corresponding author and with approval under nonexclusive use and nondisclosure contracts.

References

1. Katz Sand I, Krieger S, Farrell C, Miller AE. Diagnostic uncertainty during the transition to secondary progressive multiple sclerosis. Mult Scler 2014;20(12):1654-1657.
2. Reich DS, Lucchinetti CF, Calabresi PA. Multiple Sclerosis. N Engl J Med 2018;378(2):169-180.
3. Ciotti JR, Cross AH. Disease-modifying treatment in progressive multiple sclerosis. Curr Treat Options Neurol 2018;20(5):12.
4. LaRocca NG, Hudson LD, Rudick R, et al. The MSOAC approach to developing performance outcomes to measure and monitor multiple sclerosis disability. Mult Scler 2018;24(11):1469-1484.
5. Yousef A, Jonzzon S, Suleiman L, et al. Biosensing in multiple sclerosis. Expert Rev Med Devices 2017;14 (11):901-912.

6. Block VJ, Lizée A, Crabtree-Hartman E, et al. Continuous daily assessment of multiple sclerosis disability using remote step count monitoring. J Neurol 2017;264(2):316-326.
7. Sosnoff JJ, Sandroff BM, Pula JH, et al. Falls and physical activity in persons with multiple sclerosis. Mult Scler Int 2012;2012:315620.
8. Motl RW, Weikert M, Suh Y, et al. Accuracy of the actibelt(®) accelerometer for measuring walking speed in a controlled environment among persons with multiple sclerosis. Gait Posture 2012;35(2):192-196.
9. Sandroff BM, Motl RW, Pilutti LA, et al. Accuracy of StepWatch^{TM} and ActiGraph accelerometers for measuring steps taken among persons with multiple sclerosis. PLoS One 2014;9(4):e93511.
10. Akhbardeh A, Arjona JK, Krysko KM, et al. Novel MS vital sign: multi-sensor captures upper and lower limb dysfunction. Ann Clin Transl Neurol 2020;7(3):288-295.
11. Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. Ann Neurol 2011;69(2):292-302.
12. Lublin FD. New multiple sclerosis phenotypic classification. Eur Neurol 2014;72(Suppl 1):1-5.
13. Zeydan B, Kantarci OH. Progressive forms of multiple sclerosis: distinct entity or age-dependent phenomena. Neurol Clin 2018;36(1):163-171.
14. Faissner S, Plemel JR, Gold R, Yong VW. Progressive multiple sclerosis: from pathophysiology to therapeutic strategies. Nat Rev Drug Discov 2019;18(12):905-922.
15. Goodin DS. A questionnaire to assess neurological impairment in multiple sclerosis. Mult Scler 1998;4 (5):444-451.
16. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983;33(11):1444-1452.
17. Tanigawa M, Stein J, Park J, et al. Finger and foot tapping as alternative outcomes of upper and lower extremity function in multiple sclerosis. Mult Scler J Exp Transl Clin 2017;3(1):2055217316688930.
18. Sato S, Lim J, Miehm JD, et al. Rapid foot-tapping but not hand-tapping ability is different between relapsing-remitting and progressive multiple sclerosis. Mult Scler Relat Disord 2020;41:102031.
19. Haralick RM, Shanmugam K, Dinstein I. Textural Features for Image Classification. IEEE Trans Syst Man Cybern 1973;SMC-3(6):610-621.
20. Haralick R, Shapiro L. Computer and Robot Vision. vol. 1. Boston: Addison-Wesley, 1992.
21. Trapezoidal numerical integration - MATLAB trapz [Internet]. 2018;[cited 2018 Jun 28] Available from: https://www.mathworks.com/help/matlab/ref/trapz.html.
22. McCullagh P, Nelder J. Generalized Linear Models. New YorkChapman & Hall, 1990.
23. Molinaro AM, Simon R, Pfeiffer RM. Prediction error estimation: a comparison of resampling methods. Bioinforma Oxf Engl 2005;21(15):3301-3307.
24. Lorscheider J, Buzzard K, Jokubaitis V, et al. Defining secondary progressive multiple sclerosis. Brain 2016;139(Pt 9):2395-2405.
25. Bermel R, Waldman A, Mowry EM. Outcome measures in multiple sclerosis. Mult Scler Int 2014;2014:439375.
26. Meyer-Moock S, Feng Y-S, Maeurer M, et al. Systematic literature review and validity evaluation of the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC) in patients with multiple sclerosis. BMC Neurol 2014;14:58.
27. DeMarshall C, Goldwaser EL, Sarkar A, et al. Autoantibodies as diagnostic biomarkers for the detection and subtyping of multiple sclerosis. J Neuroimmunol 2017;309:51-57.
28. Barro C, Benkert P, Disanto G, et al. Serum neurofilament as a predictor of disease worsening and brain and spinal cord atrophy in multiple sclerosis. Brain 2018;141(8): 2382-2391.
29. Bradshaw MJ, Farrow S, Motl RW, Chitnis T. Wearable biosensors to monitor disability in multiple sclerosis. Neurol Clin Pract 2017;7(4):354-362.
30. Kappos L, Bar-Or A, Cree BAC, et al. Siponimod versus placebo in secondary progressive multiple sclerosis (EXPAND): a double-blind, randomised, phase 3 study. Lancet 2018;391(10127):1263-1273.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Supplementary Methods. Supplemental methods for extraction of Haralick features and creation of metric of within limb change in function over time

