

Original Paper

# Predicting Multiple Sclerosis Outcomes During the COVID-19 Stay-at-home Period: Observational Study Using Passively Sensed Behaviors and Digital Phenotyping

Prerna Chikersal<sup>1</sup>, MS; Shruthi Venkatesh<sup>2</sup>, BS; Karman Masown<sup>2</sup>, BS; Elizabeth Walker<sup>2</sup>, BS; Danyal Quraishi<sup>2</sup>, BS; Anind Dey<sup>3\*</sup>, PhD; Mayank Goel<sup>1\*</sup>, PhD; Zongqi Xia<sup>2\*</sup>, MD, PhD

<sup>1</sup>School of Computer Science, Carnegie Mellon University, Pittsburgh, PA, United States

<sup>2</sup>Department of Neurology, University of Pittsburgh, Pittsburgh, PA, United States

<sup>3</sup>Information School, University of Washington, Seattle, WA, United States

\*these authors contributed equally

**Corresponding Author:**

Zongqi Xia, MD, PhD  
Department of Neurology  
University of Pittsburgh  
3501 Fifth Avenue,  
BST3, Suite 7014  
Pittsburgh, PA, 15260  
United States  
Phone: 1 412 383 5377  
Email: [zxial@pitt.edu](mailto:zxial@pitt.edu)

## Abstract

**Background:** The COVID-19 pandemic has broad negative impact on the physical and mental health of people with chronic neurological disorders such as multiple sclerosis (MS).

**Objective:** We presented a machine learning approach leveraging passive sensor data from smartphones and fitness trackers of people with MS to predict their health outcomes in a natural experiment during a state-mandated stay-at-home period due to a global pandemic.

**Methods:** First, we extracted features that capture behavior changes due to the stay-at-home order. Then, we adapted and applied an existing algorithm to these behavior-change features to predict the presence of depression, high global MS symptom burden, severe fatigue, and poor sleep quality during the stay-at-home period.

**Results:** Using data collected between November 2019 and May 2020, the algorithm detected depression with an accuracy of 82.5% (65% improvement over baseline;  $F_1$ -score: 0.84), high global MS symptom burden with an accuracy of 90% (39% improvement over baseline;  $F_1$ -score: 0.93), severe fatigue with an accuracy of 75.5% (22% improvement over baseline;  $F_1$ -score: 0.80), and poor sleep quality with an accuracy of 84% (28% improvement over baseline;  $F_1$ -score: 0.84).

**Conclusions:** Our approach could help clinicians better triage patients with MS and potentially other chronic neurological disorders for interventions and aid patient self-monitoring in their own environment, particularly during extraordinarily stressful circumstances such as pandemics, which would cause drastic behavior changes.

(*JMIR Ment Health* 2022;9(8):e38495) doi: [10.2196/38495](https://doi.org/10.2196/38495)

**KEYWORDS**

mobile sensing; sensor; sensing; mobile health; mHealth; algorithm; multiple sclerosis; disability; mental health; depression; sleep; fatigue; tiredness; predict; machine learning; feature selection; neurological disorder; COVID-19; isolation; behavior change; health outcome; fitness; movement; physical activity; exercise; tracker; digital phenotyping

## Introduction

The COVID-19 pandemic and the ensuing response (eg, lockdown and social distancing) have broad negative impacts on physical and mental health worldwide [1-7]. The effect is more pronounced for people with chronic neurological diseases such as multiple sclerosis (MS) [8-10]. People with MS have a significantly higher burden of mental health comorbidities than the general population. Moreover, people with MS have a 50% lifetime prevalence of depression, 2-3 times higher than the general population [11-13]. Given its association with higher disability and mortality, depression is a major comorbidity that lowers the quality of life [11,14-20]. Further, people with MS have greater COVID-19 risk due to certain immune disease-modifying therapies as well as their physical disability, and many have experienced drastic change in their neurological care due to the pandemic [21]. Concerns for COVID-19, coupled with decreased social support and health care access during the pandemic, have contributed to even higher stress and depression in people with MS [10,22-24].

During the pandemic, digital technologies have become invaluable for supporting social interaction, health care access, and health monitoring. Digital health tools can also measure an individual's mental health profile based on passive (noninvasive) tracking. Given the complexity and heterogeneity of real-world behaviors, models that leverage different aspects of an individual's daily behaviors are necessary to accurately predict mental health status. Relevant to depression in people with MS, clinicians could use this digital passive sensing approach to potentially identify patients who require urgent health interventions.

Past research has leveraged passively generated data from personal digital devices (eg, smartphones and fitness trackers) to capture human behavior and predict health outcomes. This moment-by-moment, in situ quantification of the individual-level human phenotype using data from personal digital devices is known as digital phenotyping [25]. Previous works using passively sensed smartphone and wearable data to predict physical disability and fatigue in people with MS have been exploratory in assessing the feasibility of data collection and the preliminary association between sensed behaviors and outcomes [26-28]. However, the clinical applicability of digital phenotyping to inform clinical outcomes in people with MS in the real world has not yet been established.

Here, we present a machine learning approach leveraging data from the smartphones and fitness trackers of people with MS to predict their health outcomes during a mandatory stay-at-home period of the pandemic. Building on an existing analytical pipeline [29], we quantified behavior changes during the stay-at-home period when compared to the preceding period and used the changes to predict the presence of patient-reported outcomes of depression, neurological disability, fatigue, and poor sleep quality during the stay-at-home period. This study is different from prior studies in that it examines the clinical utility of digital phenotyping with passive sensors for predicting health outcomes during the early wave of the COVID-19 pandemic in a unique natural experiment. The study has

relevance for predicting the health outcomes of patients with chronic and complex conditions beyond MS during major stressful scenarios (eg, pandemics and natural disasters) that could considerably alter behaviors.

## Methods

### Overview

This study was part of a larger study that aimed to examine the clinical utility of passive sensors on smartphones and fitness trackers in predicting clinically relevant outcomes in people with MS. Data collection from participants in this larger study occurred between November 2019 and January 2021. Because data collection for 56 participants spanned the locally mandated stay-at-home period in response to the COVID-19 pandemic, we used this unique natural experiment to test the hypothesis whether machine learning models leveraging passive sensor data can predict the health outcomes of people with a chronic neurological disorder (ie, people with MS) during major stressful scenarios.

To briefly summarize our approach, we used data from 3 sensors in the participants' smartphones (calls, location, and screen activity) and 3 sensors in the participants' fitness trackers (heart rate, sleep, and steps) to predict patient-reported outcomes of depression, global MS symptom burden, fatigue, and sleep quality during the COVID-19 stay-at-home period. We computed behavioral features from these 6 sensors before and during the stay-at-home period and took the difference as a measure of behavior change resulting from the stay-at-home mandate. We then used changes in behavioral features to predict the outcomes.

All methods were performed in accordance with the institutional review board guideline and institutional regulation.

### Participants

The study included adults 18 years or older with a neurologist-confirmed MS diagnosis who owned a smartphone (Android or iOS) and enrolled in the Prospective Investigation of Multiple Sclerosis in the Three Rivers Region study, a clinic-based natural history study at the University of Pittsburgh Medical Center [21,30-34].

### Ethical Considerations

The institutional review boards of University of Pittsburgh (STUDY19080007) and Carnegie Mellon University (STUDY2019\_00000037) approved the study. All participants provided written informed consent.

### Study Design

The participants downloaded a mobile app to capture sensor data from their own smartphones and additionally received a Fitbit Inspire HR (Fitbit Inc) to track steps, heart rate, and sleep. Data were continuously collected from smartphone and Fitbit sensors of 56 participants during the study period (16 November 2019 to 15 May 2020, including the local stay-at-home period).

All 56 (100%) participants completed data collection for a predefined period of 12 weeks while 39 (70%) agreed to extend data collection for an additional 12 weeks (for a total of 24

weeks). Moreover, 6 (11%) participants who did not have sufficient data during the period before the stay-at-home mandate were excluded from the machine learning analysis.

### Survey Response and Patient-Reported Outcomes

All participants completed a baseline questionnaire, which queried their demographics and baseline health outcomes, on the Saturday following enrollment. During the study, the participants completed additional questionnaires, as described below, at intervals according to each questionnaire. All questionnaires for the overall study were administered via the web using the secure, web-based Research Electronic Data Capture system, including the stay-at-home period [35,36].

#### Depression

We used the Patient Health Questionnaire-9 (PHQ-9) to measure the severity of depression symptoms once every 2 weeks [37]. PHQ-9 contained 9 questions, with each answer being scored on a scale of 0 to 3. Higher scores indicated more severe depressive symptoms.

#### Global MS Symptom Burden

We used the Multiple Sclerosis Rating Scale—Revised (MSRS-R) to measure global MS symptom burden and neurological disability once every 4 weeks [38]. MSRS-R assessed 8 neurological domains (walking, upper limb function, vision, speech, swallowing, cognition, sensory, bladder, and bowel function); each domain scored as 0 to 4, with 0 indicating the absence of symptom and 4 indicating higher symptom burden and more severe disability.

#### Fatigue

We used the 5-item version of the Modified Fatigue Impact Scale (MFIS-5) to measure the impact of fatigue on cognitive, physical, and psychosocial function once every 4 weeks [39]. Each item in MFIS-5 was scored on a 5-point Likert scale from 0 (never) to 4 (almost always). Higher scores indicated more severe fatigue.

#### Sleep Quality

We used the Pittsburgh Sleep Quality Index (PSQI) to measure sleep disturbances once every 4 weeks [40]. PSQI comprised 19 individual items, with 7 component scores (each on a 0-3 scale) and 1 composite score (0 to 21, where higher scores indicate a poorer sleep quality).

For each outcome, we averaged the measures collected during the stay-at-home-period and then dichotomized the resulting outcomes using thresholds. The binary outcomes would likely have better clinical utility as they are more easily understood by patients (for self-monitoring), volunteers with limited mental health training, or even clinicians. For “Depression,” PHQ-9 scores were dichotomized as “ $\geq 5$ : presence of depression” and “ $< 5$ : absence of depression.” For “Global MS symptom burden,” MSRS-R scores were dichotomized as “ $\geq 6.4$ : higher burden” and “ $< 6.4$ : lower burden.” For “Fatigue,” MSIF-5 scores were

dichotomized as “ $\geq 8$ : high fatigue” and “ $< 8$ : low fatigue.” For “Sleep quality,” PSQI scores were dichotomized as “ $\geq 9$ : poorer sleep quality” and “ $< 9$ : better sleep quality.” The thresholds for depression and sleep quality were based on previous works [37,41]. Given the lack of consensus from the literature, we calculated the median scores of the global MS symptom burden and fatigue in a larger data set of 104 people with MS, of which the 56 (53.8%) people with MS in this paper represented a subgroup (with data collection encompassing the stay-at-home period) and used the median scores as the thresholds.

### Sensor Data Collection

Each participant installed a mobile app based on the AWARE framework [42], which provided backend and network infrastructure that unobtrusively collected from smartphones the location, screen usage (ie, when the screen status changed to on or off and locked or unlocked), and call logs (for incoming, outgoing, and missed calls). Further, participants wore a Fitbit Inspire HR, which captured the number of steps, sleep status (asleep, awake, restless, or unknown), and heart rate. Calls and screen use were event-based sensor streams, whereas location, steps, sleep, and heart rate were time series sensor streams. We sampled location coordinates at 1 sample per 10 minutes, and steps, sleep, and heart rate at 1 sample per minute.

Data from AWARE were deidentified and automatically transferred over WiFi to a study server at regular intervals. Data from the Fitbit were retrieved using the Fitbit application programming interface at the end of the data collection. Participants were asked to keep their devices charged and to always carry their phone and wear Fitbit.

To protect confidentiality, we removed identifiable information (eg, names and contact information) from survey and sensor data prior to analysis. We followed the standard practice for sensor data security.

### Mediation Analysis

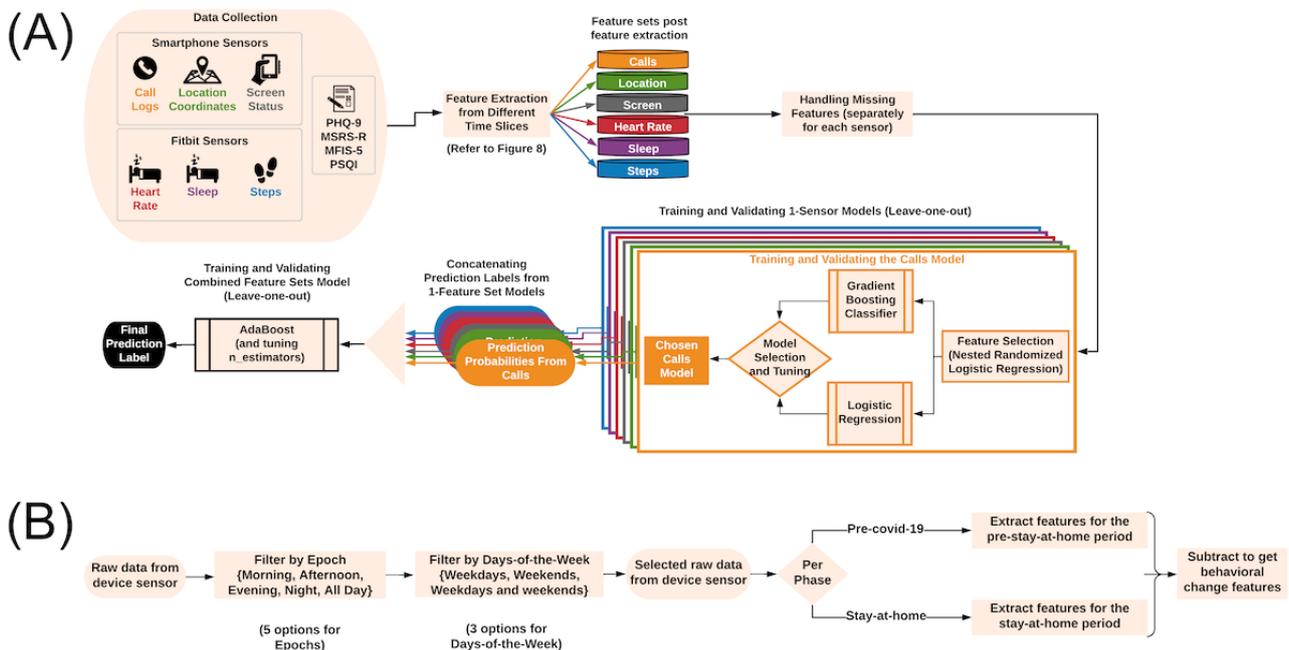
Mediation analysis was performed using the nondichotomized outcomes (ie, the average of the patient-reported outcomes collected during the stay-at-home-period). Process Macro in SPSS (IBM Corp) was used for mediation analysis [43].

### Data Processing and Machine Learning

The data processing and analysis pipeline (Figure 1) were built on our prior work [29] and involved several steps:

1. Feature extraction from sensors over time slices to identify behavior changes.
2. Handling missing features.
3. Machine learning to predict patient-reported health outcomes during the stay-at-home period:
  - Using 1-sensor models (ie, models containing features from 1 sensor).
  - Combining 1-sensor models to obtain the best model for each outcome.

**Figure 1.** Data processing and analysis pipeline. (A) Machine learning pipeline for predicting depression (Patient Health Questionnaire-9 [PHQ-9]), global MS symptom burden (Multiple Sclerosis Rating Scale—Revised [MSRS-R]), fatigue (Modified Fatigue Impact Scale-5 [MFIS-5]), and sleep quality (Pittsburgh Sleep Quality Index [PSQI]) using passive sensors from smartphones and fitness trackers. (B) For each sensor during the pre-stay-at-home period and the stay-at-home period, each feature was extracted from 15 time slices. The pre-stay-at-home features were subtracted from the stay-at-home features to obtain the behavior change features. First, raw data from the device sensor were preprocessed and then filtered by a time-of-the-day epoch and a days-of-the-week option. Features were then extracted from the selected raw data.



### Feature Extraction

We computed features from the 6 sensors of calls, heart rate, location, screen, sleep, and steps, given their potential to inform depressive symptoms [29,44-48], as well as fatigue [49], MS symptom burden such as decreased mobility [27], and sleep quality [50,51].

Location features captured mobility patterns. Steps and heart rate captured the extent of physical activities. Calls features captured communication patterns. Screen features might inform the ability for concentration [52,53] and the extent of sedentary behavior [54], despite potential caveats for people with MS and other chronic neurological disorders. Sleep features captured sleeping duration and patterns, which could indicate sleep disturbance (eg, insomnia or hypersomnia) associated with depression [55]. Please see [Multimedia Appendix 1](#) (section A.1 [29,44,46,56-59]) for details of features extracted from each sensor.

Features from the 6 sensors were extracted over a range of temporal slices ([Figure 1B](#)) preceding and during the stay-at-home period. For each period, we obtained the daily averages of these features by computing the average of the daily feature values. We computed features of behavior changes by subtracting the daily averages of features during the baseline (pre-stay-at-home) period from the stay-at-home period for the machine learning models.

### Temporal Slicing

The temporal slicing approach extracted sensor features from different time segments ([Figure 1B](#)). Past work showed that this approach can better define the relationship between a feature and depression. For example, Chow et al [60] found no

relationship between depression and the time spent at home during 4-hour time windows, but they found that people with more severe depression tended to spend more time at home between 10 AM and 6 PM. Similarly, Saeb et al [61] found that the same behavioral feature calculated over weekdays and weekends could have a very different association with depression. Here, we obtained all available data (spanning multiple days of the study) from a specific epoch or time segment of the day (all day, night [ie, 12 AM-6 AM], morning [ie, 6 AM-12 PM], afternoon [ie, 12 PM-6 PM], and evening [ie, 6 PM-12 AM]) and for specific days of the week (all days of the week, weekdays only [ie, Monday-Friday], and weekends only [ie, Saturday-Sunday]) to achieve 15 data streams or temporal slices. To extract features from each of the 15 temporal slices, we first computed daily features, averaged daily features from the pre-stay-at-home period, and averaged daily features from the stay-at-home period. We then subtracted the pre-stay-at-home feature matrix from the stay-at-home feature matrix to obtain the behavior change features. We concatenated the resulting 15 temporal slices of behavior change features to derive the final feature matrix.

### Feature Matrix

After feature extraction, each of the 6 sensors had a feature matrix, with each sample containing a participant's feature vector comprising behavior change features from 15 different temporal slices.

### Handling Missing Data

Missing sensor data can occasionally occur due to several reasons. Our approach for handling missing data is described in [Multimedia Appendix 1](#) (section A.2).

### **Machine Learning Using Nested Feature Selection**

We built machine learning models to predict dichotomized outcomes using the data set, building on a published approach [29], and validated our models using leave-5-participants-out cross-validation to minimize overfitting. The model generation process followed these steps:

1. Stable feature selection using randomized logistic regression, leveraging temporal slices.
2. Training and validating 1-sensor models for each of the 6 feature sets of calls, heart rate, location, screen, sleep, and steps.
3. Obtaining predictions from combinations of sensors by combining detection probabilities from 1-sensor models to identify the best performing model.
4. Classifying different outcomes by running the pipeline for each outcome.

### **Stable Feature Selection**

To enable stable feature selection from a vast number of behavioral features, Chikersal et al [29] proposed an approach called “nested randomized logistic regression,” which we deployed in this study. This method decomposed the feature space for each sensor by grouping features from the same time slices and performed randomized logistic regression on each of these groups. The selected features from all groups (ie, all time slices) are then concatenated to give a new and much smaller set of features. Next, we performed randomized logistic regression on this new set of features to extract the final selected features for the sensor. We performed the nested feature selection for each of the six 1-sensor models, thereby nesting the process. This method was performed in a leave-5-participants-out manner such that the model used to detect an outcome for a participant did not include that person during the feature selection process. More details about this method can be found in [Multimedia Appendix 1](#) (section A.3).

### **Training and Validating 1-Sensor Models**

For each sensor, we built a model of the selected features from that sensor to detect an outcome. We used leave-5-participants-out cross-validation to choose the parameters for that model. We trained models using the following 2 machine learning algorithms: logistic regression and gradient boosting classifier [29]. We chose the model with

the best  $F_1$ -score for a given outcome, which provides the detection probabilities for the outcome. The process is independent of other outcomes.

### **Obtaining Predictions From Combinations of Sensors**

The detection probabilities from all six 1-sensor models were concatenated into a single feature vector and given as input to an ensemble classifier (ie, AdaBoost with gradient boosting classifier as a base estimator), which then outputted the final label for the outcome. For all outcomes, only the detection probabilities of the positive label “1” were concatenated. The positive label was the “presence of depression” for “depression,” “high burden” for “global MS symptom burden,” “severe fatigue” for “fatigue,” and “poor sleep quality” for “sleep quality.” The “n\_estimators” (the maximum number of estimators at which boosting is terminated) parameter was tuned during leave-5-participants-out cross-validation to achieve the best-performing combined model.

To analyze the usefulness of each sensor, we implemented a feature ablation analysis by generating detection results for all possible combinations of 1-sensor models. For six 1-sensor models, there were 57 combinations of feature sets, as the total combinations = combinations with 2 sensors + ... + combinations with 6 sensors =

$$\sum_{r=1}^6 \binom{6}{r} = 57.$$

### **Classifying Different Outcomes**

This pipeline of training and validating six 1-sensor models and 57 combined models was run independently for each of the 4 outcomes. For each outcome, we reported the performance based on the best combination of sensors. We also reported the performance of baseline models (ie, a simple majority classifier whereby every point is assigned to whichever is in the majority in the training set) as well as models containing all 6 sensors.

## **Results**

### **Participant Characteristics**

The characteristics of the 56 participants were representative of the typical MS study (median age 43.5 years; n=48, 86% women). [Table 1](#) shows the detailed participant characteristics.

**Table 1.** Study participant characteristics.

Variable	Value
<b>Sex, n (%)</b>	
Female	48 (86)
Male	8 (14)
<b>Race, n (%)</b>	
White	51 (91)
African or African American	5 (9)
<b>Ethnicity, n (%)</b>	
Non-Hispanic or Latino	55 (98)
Hispanic or Latino	1 (2)
Age (years), median (IQR)	43.5 (37-52)
Time elapsed (years) from age of first neurological symptom onset to study participation, median (IQR)	13.0 (6.7-17.4)
PDDS <sup>a</sup> score at start of study, median (IQR)	1 (0-3)
<b>Disease-modifying treatment, n (%)</b>	
Higher efficacy	38 (68)
Standard efficacy	12 (21)
<b>Depression diagnosis, n (%)</b>	
Not diagnosed with clinical depression before study enrollment	39 (70)
Diagnosed with clinical depression before study enrollment	17 (30)
<b>Pharmacotherapy for depression, n (%)</b>	
Not taking medication for depression before study enrollment	39 (70)
Taking medication for depression before study enrollment	17 (30)
<b>Nonpharmacotherapy for depression, n (%)</b>	
Not receiving nonmedication therapy for depression before study enrollment	52 (93)
Receiving nonmedication therapy for depression before study enrollment	4 (7)
<b>Study outcomes: average measures during the stay-at-home period, median (IQR)</b>	
PHQ-9 <sup>b</sup> (depression)	3.7 (0.0-7.4)
MSRS-R <sup>c</sup> (global MS <sup>d</sup> symptom burden)	7.5 (3.4-10.3)
MFIS-5 <sup>e</sup> (fatigue)	8.0 (4.6-11.0)
PSQI <sup>f</sup> (sleep quality)	11.0 (7.8-14.3)

<sup>a</sup>PDDS: Patient Determined Disease Steps.

<sup>b</sup>PHQ-9: Patient Health Questionnaire-9.

<sup>c</sup>MSRS-R: Multiple Sclerosis Rating Scale—Revised.

<sup>d</sup>MS: multiple sclerosis.

<sup>e</sup>MFIS-5: Modified Fatigue Impact Scale-5.

<sup>f</sup>PSQI: Pittsburgh Sleep Quality Index.

## Interrelated Outcomes

The main study outcome is patient-reported depression as well as associated neurological symptom burden, fatigue, and sleep quality. We measured the Pearson correlations among the average values of the 4 outcomes during the stay-at-home period for the participants. Depression severity (PHQ-9) correlated

with the global MS symptom burden (MSRS-R), fatigue severity (MFIS-5), and sleep quality (PSQI; [Figure 2](#)).

To dissect the complex relationship among these outcomes to inform better patient monitoring and guide potentially more precise interventions, we performed mediation analysis ([Figure 3](#)). When MFIS-5 and PSQI were both included as mediators in the model (path c'), the association between MSRS-R and

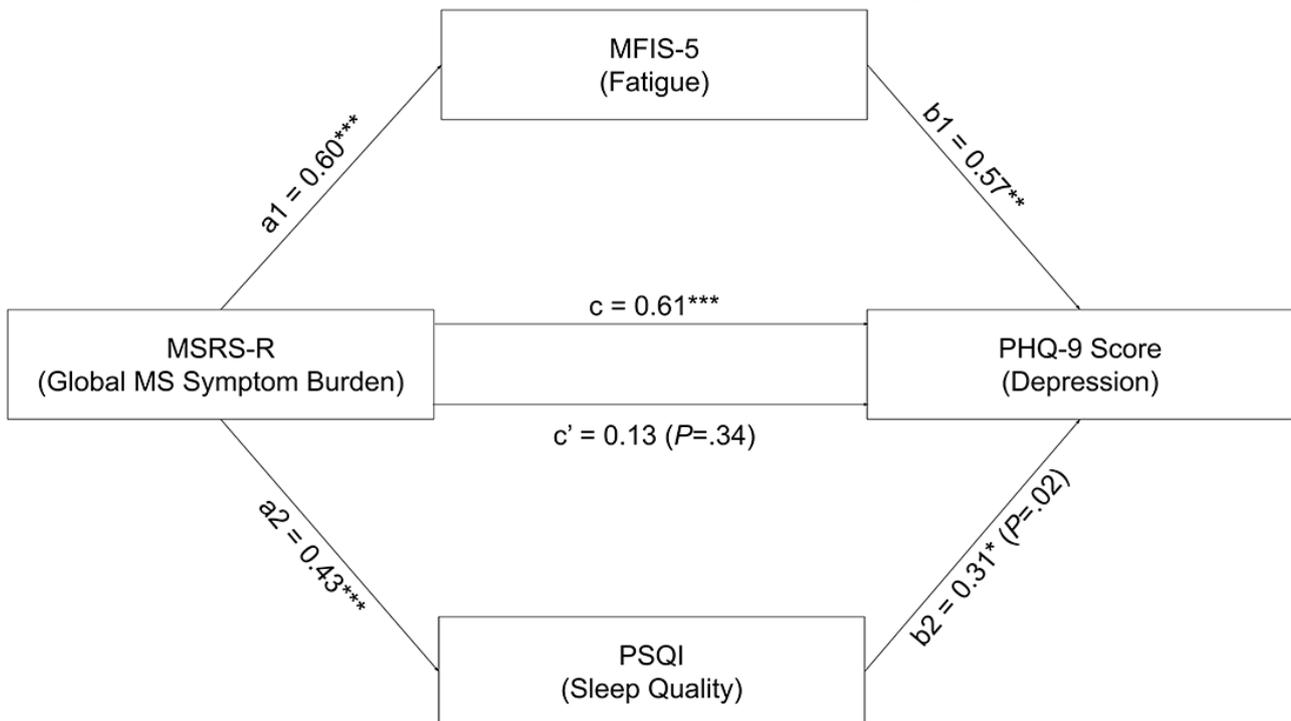
PHQ-9 was no longer significant (effect size=0.13 and the bias-corrected bootstrap confidence intervals=-0.14 and 0.40). However, the association between MSRS-R and PHQ-9 through MFIS-5 (path a1b1) remained significant (effect size=0.34 and the bias-corrected bootstrap confidence intervals=0.13-0.52).

The association between MSRS-R and PHQ-9 through PSQI (path a2b2) also remained significant (effect size=0.13 and the bias-corrected bootstrap confidence intervals=0.02-0.27). Hence, the relationship between the global MS symptom burden and depression might be mediated by both fatigue and sleep quality.

**Figure 2.** Correlations among the 4 clinically relevant patient-reported outcomes in this study. For all correlations,  $P < .001$  (N=56). MFIS-5: Modified Fatigue Impact Scale-5; MSRS-R: Multiple Sclerosis Rating Scale—Revised; PHQ-9: Patient Health Questionnaire-9; PSQI: Pittsburgh Sleep Quality Index.

	1	2	3	4
1. PHQ-9 Score (Depression)	1			
2. MSRS-R (Functional Disability)	0.602	1		
3. MFIS-5 (Fatigue)	0.713	0.73	1	
4. PSQI (Sleep Quality)	0.597	0.486	0.562	1

**Figure 3.** Parallel mediation analysis. Path model showing the effect of Multiple Sclerosis Rating Scale—Revised (MSRS-R; measuring global MS symptom burden) on the Patient Health Questionnaire-9 (PHQ-9) score (measuring depression) as mediated simultaneously by Modified Fatigue Impact Scale-5 (MFIS-5; measuring fatigue) and Pittsburgh Sleep Quality Index (PSQI; measuring sleep quality). Path c represents the effect of MSRS-R on PHQ-9 without mediators in the model. Path c' represents the effect of MSRS-R on PHQ-9 when MFIS-5 and PSQI mediators are included in the model. Paths a1b1 and a2b2 represent the effect of MSRS-R on PHQ-9 through MFIS-5 or PSQI respectively. The figure shows nonstandardized  $\beta$  regression coefficients (\* $P < .05$ , \*\* $P < .001$ , \*\*\* $P < .0001$ ) as reported by PROCESS Macro in SPSS [43]. MS: multiple sclerosis.



**Predicting Outcomes During the Stay-at-home Period**

Figure 4 shows the performance of the machine learning pipeline for predicting each of the 4 outcomes using the best sensor combinations (ie, the set of sensors that had the best performance for each outcome). Accuracy is the percentage of patients for whom the outcome label was correctly predicted.  $F_1$ -score is a metric of model performance that measures the harmonic mean of precision and recall. Precision is the positive predictive value, or the number of true positive labels divided by the number of

all positive labels (true positive + false positive). Recall is sensitivity, or the number of true positive labels divided by the number of all patients who should have the positive labels (true positive + false negative). In this study, “positive” label refers to the outcome of interest (eg, presence of depression is the positive label for depression). Figures S1 to S4 in Multimedia Appendix 1 report the performance of individual sensors and when all 6 sensors were included. Tables S1 to S4 in Multimedia Appendix 1 list the features selected by the best models for each outcome, and their corresponding coefficients.

**Figure 4.** Performance of the machine learning pipeline using all sensors and the best sensor combination for predicting each of the 4 clinically relevant outcomes in people with multiple sclerosis during a state-mandated stay-at-home period. "Accuracy (All Sensors)" and "F1 Score (All Sensors)" are the accuracy ( $\times 0.01$ ) and F1-score obtained by combining all 6 sensors. "Accuracy (Best Sensors)" and "F1 Score (Best Sensors)" are the accuracy ( $\times 0.01$ ) and F1-score obtained by the best combination of sensors. See [Multimedia Appendix 1](#) for additional performance metrics of all models. MFIS-5: Modified Fatigue Impact Scale-5; MSR-R: Multiple Sclerosis Rating Scale—Revised; PHQ-9: Patient Health Questionnaire-9; PSQI: Pittsburgh Sleep Quality Index.



### Depression

The baseline model (simple majority classifier) had an accuracy of 50.0% in predicting the presence of depression during the stay-at-home period. The model containing all sensors had an accuracy of 70% (40% improvement over the baseline). The model with the best combination of sensors (calls, heart rate, and location) had an accuracy of 82.5% (65% improvement over the baseline).

### Global MS Symptom Burden

The baseline model had an accuracy of 64.7% in predicting high global MS symptom burden (versus "low burden") during the stay-at-home period. The model containing all sensors had an accuracy of 76.7% (18.5% improvement over the baseline). The model with the best combination of sensors (calls, heart rate, location, and screen) had an accuracy of 90% (39% improvement over the baseline).

### Fatigue

The baseline model had an accuracy of 61.8% in predicting severe fatigue (versus "mild fatigue") during the stay-at-home period. The model containing all sensors had an accuracy of 71.7% (16% improvement over the baseline). The model with the best combination of sensors (calls, heart rate, and location) had an accuracy of 75.5% (22% improvement over the baseline).

### Sleep Quality

The baseline model had an accuracy of 65.7% in predicting poor sleep quality (ie, "poor sleep quality" versus "better sleep quality") during the stay-at-home period. The model containing all sensors had an accuracy of 70.2% (7% improvement over the baseline). The model with the best combination of sensors (location and screen) had an accuracy of 84% (28% improvement over the baseline).

## Discussion

### Principal Findings

In this unique natural experiment conducted during the early wave of the COVID-19 pandemic, we reported the clinical utility of digital phenotyping for predicting clinically relevant outcomes

for people with MS. Using only passively sensed data, our machine-learning models predicted the presence of depression, high global MS symptom burden, severe fatigue, and poor sleep quality during the stay-at-home period with potentially clinically actionable performance.

The best models outperformed not only baseline models (simple majority classifier) but also models containing all sensors. The best sensor combinations for predicting depression and fatigue were the same (ie, calls, heart rate, and location), while these sensors were also included in the best sensor combination for predicting global MS symptom burden (ie, calls, heart rate, location, and screen). Comparably, the best sensor combination for sleep quality (ie, location and screen) had the smallest overlap with the sensor combinations for the other three outcomes. This observation was consistent with the finding that depression, fatigue, and global MS symptom burden were better correlated among themselves than with sleep quality ([Figure 2](#)). We also looked at the feature coefficients of the features selected by the best models ([Multimedia Appendix 1](#), section B.2). Examples of the best features of changed behavior selected by the best model for predicting depression (ie, features with the highest absolute coefficients) include increase in number of incoming calls during evenings on weekdays, decrease in average heart rate when the person is at rest or has low activity (outside exercise heart rate zones) during evenings on weekends, and increase in regularity in movement patterns in 24-hour periods with respect to nights on weekends.

Our findings built on a small body of prior work that explored the feasibility of passive sensing in people with MS and preliminary correlations between passively sensed behaviors and MS outcomes. For example, Newland et al [26] explored real-time depth sensors at home to identify gait disturbance and falls in 21 patients with MS. Other studies reported correlations between passively sensed physical activity and disability worsening in people with MS [27,62,63]. Chitnis et al [28] examined the gait, mobility, and sleep of 25 people with MS over 8 weeks using sensors mounted on their wrist, ankle, and sternum, and reported correlations among gait-related features (eg, turn angle and maximum angular velocity), sleep and activity, and disability outcomes.

Previous work on predicting health outcomes for people with MS using passively sensed behaviors is scarce. Tong et al [49] used passively sensed sleep and activity data collected from 198 people with MS over 6 months to predict fatigue severity and overall health scores, achieving good performance in line with acceptable instrument errors. To our knowledge, our study is the first to use passively sensed behavior changes to predict multiple interrelated clinically relevant health outcomes in MS, including depression, disability, fatigue, and sleep quality. While several studies used passively sensed data from the general population to report behavior changes during the COVID-19 pandemic [64-67], our study provides the first real-world evidence of potential clinical utility of passively sensed behavior changes to predict health outcomes during the unique stay-at-home period in a population with a chronic neurological disorder and complex health needs. From a methodological standpoint, the application of behavioral features computed over temporal slices to predict depression and other health outcomes in people with MS is novel. Our approach of using change in features between the period preceding the stay-at-home and stay-at-home periods to predict outcomes during the stay-at-home period is also novel. Finally, we included new heart rate features that can be computed using data from the Fitbit application programming interface.

Our approach has potential clinical utility, particularly during major stressful events (beyond COVID-19) that worsen health outcomes and limit health care access. For instance, predictive models built using our approach could help patients self-monitor their health when access to in-person clinical care becomes suddenly limited and could encourage patients (or their caregivers) to actively seek medical attention sooner when the models predict adverse outcomes. Further, our models could help clinicians better monitor at-risk patients and make triage decisions for patients who require prioritization for interventions (eg, medication and counseling), particularly in the setting of suddenly limited health care access and scarce resources.

### Limitations

Our study has 2 limitations. First, the COVID-19 pandemic started during our data collection for an ongoing larger study of people with MS. While it provided a unique opportunity to conduct a natural experiment to assess the utility of digital phenotyping to predict health outcomes in people with MS

during the highly unusual stay-at-home period, we had a modest sample size of participants who happened to have sufficient sensor data collected both just before the sudden issue of the stay-at-home order and during the stay-at-home period. We also had limited ability to seek external replication of the drastic behavior changes during the early stage of the pandemic since the stay-at-home order was lifted and has not been reinstated. To reduce the chance of overfitting and improve the validity of the findings, we used leave-5-participants-out cross-validation, such that in each fold, the participants used for training and testing were different. Our approach performed well for not only 1 outcome but all 4 clinically relevant outcomes pertaining to mental health and neurological disability in people with MS. We have reasonable confidence because of the consistently good model performance across all 5 folds and the consistently robust predictions for all 4 outcomes. We are not aware of other published studies with data from before and during the stay-at-home orders, particularly involving patient population with chronic neurological disorders such as MS who are at heightened risk for adverse health outcomes resulting from social isolation, reduced support, and limited health care access. Given the uniqueness of the data set, we believe the findings are clinically relevant despite the relatively modest sample size. Second, the study used patient-reported health outcomes. Given the restriction of in-person clinical visits during the stay-at-home period, rater-performed examination was not feasible. Importantly, these patient-reported outcomes are all validated for people with MS, highly correlated with rater-determined measures, interrelated among themselves, and clinically relevant.

In summary, we reported the potential clinical utility of digital phenotyping in predicting subsequent health outcomes in people with MS during a COVID-19 stay-at-home period. Specifically, we predicted the presence of depression, high global MS symptom burden, severe fatigue, and poor sleep quality in people with MS during the stay-at-home period using passively sensed behavior changes measured by smartphone and wearable fitness tracker. The predictive models achieved potentially clinically actionable performance for all 4 outcomes. This study paved the way for future replication studies during major stressful events and has implications for future patient self-monitoring and clinician screening for urgent interventions in MS and other complex chronic diseases.

### Acknowledgments

We would like to thank our undergraduate research assistants: Man Jun (John) Han, Dong Yun Lee, Kasey Park, Phoebe Soong, and Christine Wu, for helping us monitor participant compliance throughout the data collection process. We would also like to thank Yiyi Ren for helping develop the app used for data collection. We would also like to thank the research participants and their treating clinicians. The study is funded by the Department of Defense (CDMRP MS190178).

### Authors' Contributions

PC designed and conceptualized the study; analyzed and interpreted data; and drafted and revised the manuscript for intellectual content. SV, KM, EW, and DQ played a major role in the data acquisition. AD and MG designed and conceptualized the study; interpreted the data; and drafted and revised the manuscript for intellectual content. ZX designed and conceptualized the study and had a major role in the data acquisition, data interpretation, drafting, and revision of the manuscript for intellectual content.

## Conflicts of Interest

None declared.

## Multimedia Appendix 1

Supplementary material.

[[DOCX File , 18995 KB-Multimedia Appendix 1](#)]

## References

1. Czeisler M, Lane RI, Petrosky E, Wiley JF, Christensen A, Njai R, et al. Mental health, substance use, and suicidal ideation during the COVID-19 pandemic - United States, June 24-30, 2020. *MMWR Morb Mortal Wkly Rep* 2020 Aug 14;69(32):1049-1057 [FREE Full text] [doi: [10.15585/mmwr.mm6932a1](https://doi.org/10.15585/mmwr.mm6932a1)] [Medline: [32790653](https://pubmed.ncbi.nlm.nih.gov/32790653/)]
2. Panchal N, Kamal R, Cox C, Garfield R. The implications of covid-19 for mental health and substance use. Kaiser family foundation. 2020. URL: <http://medfam.facmed.unam.mx/wp-content/uploads/2021/05/implicaciones-de-COVID-EN-LA-SALUD-MENTAL.pdf> [accessed 2022-07-27]
3. Ettman CK, Abdalla SM, Cohen GH, Sampson L, Vivier PM, Galea S. Prevalence of depression symptoms in US adults before and during the COVID-19 pandemic. *JAMA Netw Open* 2020 Sep 01;3(9):e2019686 [FREE Full text] [doi: [10.1001/jamanetworkopen.2020.19686](https://doi.org/10.1001/jamanetworkopen.2020.19686)] [Medline: [32876685](https://pubmed.ncbi.nlm.nih.gov/32876685/)]
4. Twenge JM, Joiner TE. U.S. Census Bureau-assessed prevalence of anxiety and depressive symptoms in 2019 and during the 2020 COVID-19 pandemic. *Depress Anxiety* 2020 Oct;37(10):954-956 [FREE Full text] [doi: [10.1002/da.23077](https://doi.org/10.1002/da.23077)] [Medline: [32667081](https://pubmed.ncbi.nlm.nih.gov/32667081/)]
5. Kujawa A, Green H, Compas BE, Dickey L, Pegg S. Exposure to COVID-19 pandemic stress: Associations with depression and anxiety in emerging adults in the United States. *Depress Anxiety* 2020 Dec 10;37(12):1280-1288. [doi: [10.1002/da.23109](https://doi.org/10.1002/da.23109)] [Medline: [33169481](https://pubmed.ncbi.nlm.nih.gov/33169481/)]
6. O'Connor RC, Wetherall K, Cleare S, McClelland H, Melson AJ, Niedzwiedz CL, et al. Mental health and well-being during the COVID-19 pandemic: longitudinal analyses of adults in the UK COVID-19 mental health & wellbeing study. *Br J Psychiatry* 2021 Jun;218(6):326-333 [FREE Full text] [doi: [10.1192/bjp.2020.212](https://doi.org/10.1192/bjp.2020.212)] [Medline: [33081860](https://pubmed.ncbi.nlm.nih.gov/33081860/)]
7. Lebel C, MacKinnon A, Bagshawe M, Tomfohr-Madsen L, Giesbrecht G. Elevated depression and anxiety symptoms among pregnant individuals during the COVID-19 pandemic. *J Affect Disord* 2020 Dec 01;277:5-13 [FREE Full text] [doi: [10.1016/j.jad.2020.07.126](https://doi.org/10.1016/j.jad.2020.07.126)] [Medline: [32777604](https://pubmed.ncbi.nlm.nih.gov/32777604/)]
8. Motolese F, Rossi M, Albergo G, Stelitano D, Villanova M, Di Lazzaro V, et al. The psychological impact of COVID-19 pandemic on people with multiple sclerosis. *Front Neurol* 2020 Oct 30;11:580507 [FREE Full text] [doi: [10.3389/fneur.2020.580507](https://doi.org/10.3389/fneur.2020.580507)] [Medline: [33193033](https://pubmed.ncbi.nlm.nih.gov/33193033/)]
9. Zanghi A, D'Amico E, Luca M, Ciaarella M, Basile L, Patti F. Mental health status of relapsing-remitting multiple sclerosis Italian patients returning to work soon after the easing of lockdown during COVID-19 pandemic: A monocentric experience. *Mult Scler Relat Disord* 2020 Nov;46:102561 [FREE Full text] [doi: [10.1016/j.msard.2020.102561](https://doi.org/10.1016/j.msard.2020.102561)] [Medline: [33045494](https://pubmed.ncbi.nlm.nih.gov/33045494/)]
10. Broche-Pérez Y, Jiménez-Morales RM, Monasterio-Ramos LO, Vázquez-Gómez LA, Fernández-Fleites Z. Fear of COVID-19, problems accessing medical appointments, and subjective experience of disease progression, predict anxiety and depression reactions in patients with Multiple Sclerosis. *Mult Scler Relat Disord* 2021 Aug;53:103070 [FREE Full text] [doi: [10.1016/j.msard.2021.103070](https://doi.org/10.1016/j.msard.2021.103070)] [Medline: [34119745](https://pubmed.ncbi.nlm.nih.gov/34119745/)]
11. Patten SB, Marrie RA, Carta MG. Depression in multiple sclerosis. *Int Rev Psychiatry* 2017 Oct 06;29(5):463-472. [doi: [10.1080/09540261.2017.1322555](https://doi.org/10.1080/09540261.2017.1322555)] [Medline: [28681616](https://pubmed.ncbi.nlm.nih.gov/28681616/)]
12. Chan CK, Tian F, Pimentel Maldonado D, Mowry EM, Fitzgerald KC. Depression in multiple sclerosis across the adult lifespan. *Mult Scler* 2021 Oct 14;27(11):1771-1780 [FREE Full text] [doi: [10.1177/1352458520979304](https://doi.org/10.1177/1352458520979304)] [Medline: [33307963](https://pubmed.ncbi.nlm.nih.gov/33307963/)]
13. Solaro C, Gamberini G, Masuccio FG. Depression in multiple sclerosis: epidemiology, aetiology, diagnosis and treatment. *CNS Drugs* 2018 Feb 7;32(2):117-133. [doi: [10.1007/s40263-018-0489-5](https://doi.org/10.1007/s40263-018-0489-5)] [Medline: [29417493](https://pubmed.ncbi.nlm.nih.gov/29417493/)]
14. Siegert RJ, Abernethy DA. Depression in multiple sclerosis: a review. *J Neurol Neurosurg Psychiatry* 2005 Apr 01;76(4):469-475 [FREE Full text] [doi: [10.1136/jnnp.2004.054635](https://doi.org/10.1136/jnnp.2004.054635)] [Medline: [15774430](https://pubmed.ncbi.nlm.nih.gov/15774430/)]
15. Feinstein A, Magalhaes S, Richard J, Audet B, Moore C. The link between multiple sclerosis and depression. *Nat Rev Neurol* 2014 Sep 12;10(9):507-517. [doi: [10.1038/nrneuro.2014.139](https://doi.org/10.1038/nrneuro.2014.139)] [Medline: [25112509](https://pubmed.ncbi.nlm.nih.gov/25112509/)]
16. Zhang Y, Taylor BV, Simpson S, Blizzard L, Campbell JA, Palmer AJ, et al. Feelings of depression, pain and walking difficulties have the largest impact on the quality of life of people with multiple sclerosis, irrespective of clinical phenotype. *Mult Scler* 2021 Jul 14;27(8):1262-1275. [doi: [10.1177/1352458520958369](https://doi.org/10.1177/1352458520958369)] [Medline: [32924841](https://pubmed.ncbi.nlm.nih.gov/32924841/)]
17. Diamond BJ, Johnson SK, Kaufman M, Graves L. Relationships between information processing, depression, fatigue and cognition in multiple sclerosis. *Arch Clin Neuropsychol* 2008 Mar;23(2):189-199. [doi: [10.1016/j.acn.2007.10.002](https://doi.org/10.1016/j.acn.2007.10.002)] [Medline: [18053682](https://pubmed.ncbi.nlm.nih.gov/18053682/)]
18. Ford H, Trigwell P, Johnson M. The nature of fatigue in multiple sclerosis. *Journal of Psychosomatic Research* 1998 Jul;45(1):33-38. [doi: [10.1016/s0022-3999\(98\)00004-x](https://doi.org/10.1016/s0022-3999(98)00004-x)]

19. Bakshi R, Shaikh ZA, Miletich RS, Czarnecki D, Dmochowski J, Henschel K, et al. Fatigue in multiple sclerosis and its relationship to depression and neurologic disability. *Mult Scler* 2000 Jun 02;6(3):181-185. [doi: [10.1177/135245850000600308](https://doi.org/10.1177/135245850000600308)] [Medline: [10871830](https://pubmed.ncbi.nlm.nih.gov/10871830/)]
20. Strober LB, Arnett PA. An examination of four models predicting fatigue in multiple sclerosis. *Arch Clin Neuropsychol* 2005 Jul;20(5):631-646. [doi: [10.1016/j.acn.2005.04.002](https://doi.org/10.1016/j.acn.2005.04.002)] [Medline: [15894455](https://pubmed.ncbi.nlm.nih.gov/15894455/)]
21. Levin SN, Venkatesh S, Nelson KE, Li Y, Aguerre I, Zhu W, Multiple Sclerosis Resilience to COVID-19 (MSReCOV) Collaborative. Manifestations and impact of the COVID-19 pandemic in neuroinflammatory diseases. *Ann Clin Transl Neurol* 2021 Apr 22;8(4):918-928 [FREE Full text] [doi: [10.1002/acn3.51314](https://doi.org/10.1002/acn3.51314)] [Medline: [33616290](https://pubmed.ncbi.nlm.nih.gov/33616290/)]
22. Vogel AC, Schmidt H, Loud S, McBurney R, Mateen FJ. Impact of the COVID-19 pandemic on the health care of >1,000 People living with multiple sclerosis: A cross-sectional study. *Mult Scler Relat Disord* 2020 Nov;46:102512 [FREE Full text] [doi: [10.1016/j.msard.2020.102512](https://doi.org/10.1016/j.msard.2020.102512)] [Medline: [32977074](https://pubmed.ncbi.nlm.nih.gov/32977074/)]
23. Manacorda T, Bandiera P, Terzuoli F, Ponzio M, Brichetto G, Zaratin P, et al. Impact of the COVID-19 pandemic on persons with multiple sclerosis: Early findings from a survey on disruptions in care and self-reported outcomes. *J Health Serv Res Policy* 2021 Jul 18;26(3):189-197 [FREE Full text] [doi: [10.1177/1355819620975069](https://doi.org/10.1177/1355819620975069)] [Medline: [33337256](https://pubmed.ncbi.nlm.nih.gov/33337256/)]
24. Levit E, Cohen I, Dahl M, Edwards K, Weinstock-Guttman B, Ishikawa T, Multiple Sclerosis Resilience to COVID-19 (MSReCOV) Collaborative. Worsening physical functioning in patients with neuroinflammatory disease during the COVID-19 pandemic. *Mult Scler Relat Disord* 2022 Feb;58:103482 [FREE Full text] [doi: [10.1016/j.msard.2021.103482](https://doi.org/10.1016/j.msard.2021.103482)] [Medline: [35016114](https://pubmed.ncbi.nlm.nih.gov/35016114/)]
25. Huckvale K, Venkatesh S, Christensen H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *NPJ Digit Med* 2019 Sep 6;2(1):88 [FREE Full text] [doi: [10.1038/s41746-019-0166-1](https://doi.org/10.1038/s41746-019-0166-1)] [Medline: [31508498](https://pubmed.ncbi.nlm.nih.gov/31508498/)]
26. Newland P, Wagner JM, Salter A, Thomas FP, Skubic M, Rantz M. Exploring the feasibility and acceptability of sensor monitoring of gait and falls in the homes of persons with multiple sclerosis. *Gait Posture* 2016 Sep;49:277-282. [doi: [10.1016/j.gaitpost.2016.07.005](https://doi.org/10.1016/j.gaitpost.2016.07.005)] [Medline: [27474948](https://pubmed.ncbi.nlm.nih.gov/27474948/)]
27. Shammass L, Zentek T, von Haaren B, Schlesinger S, Hey S, Rashid A. Home-based system for physical activity monitoring in patients with multiple sclerosis (Pilot study). *Biomed Eng Online* 2014 Feb 06;13:10 [FREE Full text] [doi: [10.1186/1475-925X-13-10](https://doi.org/10.1186/1475-925X-13-10)] [Medline: [24502230](https://pubmed.ncbi.nlm.nih.gov/24502230/)]
28. Chitnis T, Glanz BI, Gonzalez C, Healy BC, Saraceno TJ, Sattarnezhad N, et al. Quantifying neurologic disease using biosensor measurements in-clinic and in free-living settings in multiple sclerosis. *NPJ Digit Med* 2019 Dec 11;2(1):123 [FREE Full text] [doi: [10.1038/s41746-019-0197-7](https://doi.org/10.1038/s41746-019-0197-7)] [Medline: [31840094](https://pubmed.ncbi.nlm.nih.gov/31840094/)]
29. Chikersal P, Doryab A, Tumminia M, Villalba DK, Dutcher JM, Liu X, et al. Detecting depression and predicting its onset using longitudinal symptoms captured by passive sensing. *ACM Trans. Comput.-Hum. Interact* 2021 Feb 28;28(1):1-41. [doi: [10.1145/3422821](https://doi.org/10.1145/3422821)]
30. Levin SN, Riley CS, Dhand A, White CC, Venkatesh S, Boehm B, et al. Association of social network structure and physical function in patients with multiple sclerosis. *Neurology* 2020 Aug 07;95(11):e1565-e1574. [doi: [10.1212/wnl.00000000000010460](https://doi.org/10.1212/wnl.00000000000010460)]
31. Mani A, Santini T, Puppala R, Dahl M, Venkatesh S, Walker E, et al. Applying deep learning to accelerated clinical brain magnetic resonance imaging for multiple sclerosis. *Front Neurol* 2021 Sep 27;12:685276 [FREE Full text] [doi: [10.3389/fneur.2021.685276](https://doi.org/10.3389/fneur.2021.685276)] [Medline: [34646227](https://pubmed.ncbi.nlm.nih.gov/34646227/)]
32. Boorgu DS, Venkatesh S, Lakhani CM, Walker E, Aguerre IM, Riley C, et al. The impact of socioeconomic status on subsequent neurological outcomes in multiple sclerosis. *Mult Scler Relat Disord* 2022 Jun;65:103994. [doi: [10.1016/j.msard.2022.103994](https://doi.org/10.1016/j.msard.2022.103994)]
33. Kever A, Walker ELS, Riley CS, Heyman RA, Xia Z, Leavitt VM. Association of personality traits with physical function, cognition, and mood in multiple sclerosis. *Mult Scler Relat Disord* 2022 Feb;59:103648. [doi: [10.1016/j.msard.2022.103648](https://doi.org/10.1016/j.msard.2022.103648)]
34. Epstein S, Xia Z, Lee AJ, Dahk M, Edwards K, Levit E, Multiple Sclerosis Resilience to COVID-19 (MSReCOV) Collaborative. . Vaccination against SARS-CoV-2 in neuroinflammatory disease: early safety/tolerability data. *Mult Scler Relat Disord* 2022 Jan;57:103433. [doi: [10.1016/j.msard.2021.103433](https://doi.org/10.1016/j.msard.2021.103433)]
35. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, REDCap Consortium. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019 Jul;95:103208 [FREE Full text] [doi: [10.1016/j.jbi.2019.103208](https://doi.org/10.1016/j.jbi.2019.103208)] [Medline: [31078660](https://pubmed.ncbi.nlm.nih.gov/31078660/)]
36. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009 Apr;42(2):377-381 [FREE Full text] [doi: [10.1016/j.jbi.2008.08.010](https://doi.org/10.1016/j.jbi.2008.08.010)] [Medline: [18929686](https://pubmed.ncbi.nlm.nih.gov/18929686/)]
37. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
38. Wicks P, Vaughan TE, Massagli MP. The multiple sclerosis rating scale, revised (MSRS-R): development, refinement, and psychometric validation using an online community. *Health Qual Life Outcomes* 2012 Jun 18;10:70 [FREE Full text] [doi: [10.1186/1477-7525-10-70](https://doi.org/10.1186/1477-7525-10-70)] [Medline: [22709981](https://pubmed.ncbi.nlm.nih.gov/22709981/)]

39. Meca-Lallana V, Brañas-Pampillón M, Higuera Y, Candelieri-Merlicco A, Aladro-Benito Y, Rodríguez-De la Fuente O, et al. Assessing fatigue in multiple sclerosis: Psychometric properties of the five-item Modified Fatigue Impact Scale (MFIS-5). *Mult Scler J Exp Transl Clin* 2019 Nov 09;5(4):2055217319887987 [FREE Full text] [doi: [10.1177/2055217319887987](https://doi.org/10.1177/2055217319887987)] [Medline: [31741743](https://pubmed.ncbi.nlm.nih.gov/31741743/)]
40. Buysse DJ, Reynolds CIII, Monk TH, Hoch CC, Yeager AL, Kupfer DJ. Quantification of subjective sleep quality in healthy elderly men and women using the Pittsburgh sleep quality index (PSQI). *Sleep* 1991;14(4):331-338. [doi: [10.1093/sleep/14.4.331](https://doi.org/10.1093/sleep/14.4.331)]
41. Fictenberg NL, Putnam SH, Mann NR, Zafonte RD, Millard AE. Insomnia screening in postacute traumatic brain injury: utility and validity of the Pittsburgh Sleep Quality Index. *Am J Phys Med Rehabil* 2001 May;80(5):339-345. [doi: [10.1097/00002060-200105000-00003](https://doi.org/10.1097/00002060-200105000-00003)] [Medline: [11327555](https://pubmed.ncbi.nlm.nih.gov/11327555/)]
42. Ferreira D, Kostakos V, Dey AK. AWARE: mobile context instrumentation framework. *Front. ICT* 2015 Apr 20;2:1-9. [doi: [10.3389/fict.2015.00006](https://doi.org/10.3389/fict.2015.00006)]
43. Hayes A. Introduction to Mediation, Moderation, and Conditional Process Analysis: A Regression-Based Approach. New York, US: Guilford Press; 2017.
44. Saeb S, Zhang M, Kwasny MM, Karr CJ, Kording K, Mohr DC. The relationship between clinical, momentary, and sensor-based assessment of depression. *Int Conf Pervasive Comput Technol Healthc* 2015 Aug;2015:1-10 [FREE Full text] [doi: [10.4108/icst.pervasivehealth.2015.259034](https://doi.org/10.4108/icst.pervasivehealth.2015.259034)] [Medline: [26640739](https://pubmed.ncbi.nlm.nih.gov/26640739/)]
45. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. Studentlife: assessing mental health, academic performance and behavioral trends of college students using smartphones. 2014 Presented at: ACM international joint conference on pervasive and ubiquitous computing; September 13-17, 2014; Seattle, Washington. [doi: [10.1145/2632048.2632054](https://doi.org/10.1145/2632048.2632054)]
46. Canzian L, Musolesi M. Trajectories of depression: unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. 2015 Presented at: ACM international joint conference on pervasive and ubiquitous computing; September 7-11, 2015; Osaka, Japan p. 1293-1304. [doi: [10.1145/2750858.2805845](https://doi.org/10.1145/2750858.2805845)]
47. Xu X, Chikersal P, Dutcher JM, Sefidgar YS, Seo W, Tumminia MJ, et al. Leveraging collaborative-filtering for personalized behavior modeling. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol* 2021 Mar 19;5(1):1-27. [doi: [10.1145/3448107](https://doi.org/10.1145/3448107)]
48. Xu X, Chikersal P, Doryab A, Villalba DK, Dutcher JM, Tumminia MJ, et al. Leveraging routine behavior and contextually-filtered features for depression detection among college students. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol* 2019 Sep 09;3(3):1-33. [doi: [10.1145/3351274](https://doi.org/10.1145/3351274)]
49. Tong C, Craner M, Vegreville M, Lane ND. Tracking fatigue and health state in multiple sclerosis patients using connected wellness devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol* 2019 Sep 09;3(3):1-19. [doi: [10.1145/3351264](https://doi.org/10.1145/3351264)]
50. Min JK, Doryab A, Wiese J, Amini S, Zimmerman J, Hong JI. Toss 'n' turn: smartphone as sleep and sleep quality detector. 2014 Presented at: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; April 26, 2014; Toronto, Canada p. 477-486. [doi: [10.1145/2556288.2557220](https://doi.org/10.1145/2556288.2557220)]
51. Sano A, Phillips AJ, Yu AZ, McHill AW, Taylor S, Jaques N, et al. Recognizing academic performance, sleep quality, stress level, and mental health using personality traits, wearable sensors and mobile phones. 2015 Presented at: IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN); October 19, 2015; Cambridge, MA, USA p. 1-6.
52. Demirci K, Akgönül M, Akpınar A. Relationship of smartphone use severity with sleep quality, depression, and anxiety in university students. *J Behav Addict* 2015 Jun;4(2):85-92 [FREE Full text] [doi: [10.1556/2006.4.2015.010](https://doi.org/10.1556/2006.4.2015.010)] [Medline: [26132913](https://pubmed.ncbi.nlm.nih.gov/26132913/)]
53. Kwon M, Lee J, Won W, Park J, Min J, Hahn C, et al. Development and validation of a smartphone addiction scale (SAS). *PLoS One* 2013;8(2):e56936 [FREE Full text] [doi: [10.1371/journal.pone.0056936](https://doi.org/10.1371/journal.pone.0056936)] [Medline: [23468893](https://pubmed.ncbi.nlm.nih.gov/23468893/)]
54. Costigan SA, Barnett L, Plotnikoff RC, Lubans DR. The health indicators associated with screen-based sedentary behavior among adolescent girls: a systematic review. *J Adolesc Health* 2013 Apr;52(4):382-392. [doi: [10.1016/j.jadohealth.2012.07.018](https://doi.org/10.1016/j.jadohealth.2012.07.018)] [Medline: [23299000](https://pubmed.ncbi.nlm.nih.gov/23299000/)]
55. Nutt D, Wilson S, Paterson L. Sleep disorders as core symptoms of depression. *Dialogues in Clinical Neuroscience* 2022 Apr 01;10(3):329-336. [doi: [10.31887/dcms.2008.10.3/dnutt](https://doi.org/10.31887/dcms.2008.10.3/dnutt)]
56. Press WH, Rybicki GB. Fast algorithm for spectral analysis of unevenly sampled data. *The Astrophysical Journal*. URL: <https://adsabs.harvard.edu/pdf/1989ApJ...338..277P> [accessed 2022-07-27]
57. Mantua J, Gravel N, Spencer RMC. Reliability of sleep measures from four personal health monitoring devices compared to research-based actigraphy and polysomnography. *Sensors (Basel)* 2016 May 05;16(5):646 [FREE Full text] [doi: [10.3390/s16050646](https://doi.org/10.3390/s16050646)] [Medline: [27164110](https://pubmed.ncbi.nlm.nih.gov/27164110/)]
58. Cook JD, Prairie ML, Plante DT. Utility of the Fitbit Flex to evaluate sleep in major depressive disorder: A comparison against polysomnography and wrist-worn actigraphy. *J Affect Disord* 2017 Aug 01;217:299-305 [FREE Full text] [doi: [10.1016/j.jad.2017.04.030](https://doi.org/10.1016/j.jad.2017.04.030)] [Medline: [28448949](https://pubmed.ncbi.nlm.nih.gov/28448949/)]
59. de Zambotti M, Goldstone A, Claudatos S, Colrain IM, Baker FC. A validation study of Fitbit Charge 2™ compared with polysomnography in adults. *Chronobiol Int* 2018 Apr;35(4):465-476. [doi: [10.1080/07420528.2017.1413578](https://doi.org/10.1080/07420528.2017.1413578)] [Medline: [29235907](https://pubmed.ncbi.nlm.nih.gov/29235907/)]

60. Chow PI, Fua K, Huang Y, Bonelli W, Xiong H, Barnes LE, et al. Using mobile sensing to test clinical models of depression, social anxiety, state affect, and social isolation among college students. *J Med Internet Res* 2017 Mar 03;19(3):e62 [[FREE Full text](#)] [doi: [10.2196/jmir.6820](https://doi.org/10.2196/jmir.6820)] [Medline: [28258049](https://pubmed.ncbi.nlm.nih.gov/28258049/)]
61. Saeb S, Lattie EG, Schueller SM, Kording KP, Mohr DC. The relationship between mobile phone location sensor data and depressive symptom severity. *PeerJ* 2016;4:e2537 [[FREE Full text](#)] [doi: [10.7717/peerj.2537](https://doi.org/10.7717/peerj.2537)] [Medline: [28344895](https://pubmed.ncbi.nlm.nih.gov/28344895/)]
62. Block VJ, Bove R, Zhao C, Garcha P, Graves J, Romeo AR, et al. Association of continuous assessment of step Count by remote monitoring with disability progression among adults with multiple sclerosis. *JAMA Netw Open* 2019 Mar 01;2(3):e190570 [[FREE Full text](#)] [doi: [10.1001/jamanetworkopen.2019.0570](https://doi.org/10.1001/jamanetworkopen.2019.0570)] [Medline: [30874777](https://pubmed.ncbi.nlm.nih.gov/30874777/)]
63. Stuart CM, Varatharaj A, Domjan J, Philip S, Galea I, SIMS study group. Physical activity monitoring to assess disability progression in multiple sclerosis. *Mult Scler J Exp Transl Clin* 2020 Dec 07;6(4):2055217320975185 [[FREE Full text](#)] [doi: [10.1177/2055217320975185](https://doi.org/10.1177/2055217320975185)] [Medline: [33343919](https://pubmed.ncbi.nlm.nih.gov/33343919/)]
64. Sun S, Folarin AA, Ranjan Y, Rashid Z, Conde P, Stewart C, RADAR-CNS Consortium. Using smartphones and wearable devices to monitor behavioral changes during COVID-19. *J Med Internet Res* 2020 Sep 25;22(9):e19992 [[FREE Full text](#)] [doi: [10.2196/19992](https://doi.org/10.2196/19992)] [Medline: [32877352](https://pubmed.ncbi.nlm.nih.gov/32877352/)]
65. Ong J, Lau T, Massar SAA, Chong ZT, Ng BKL, Koek D, et al. COVID-19-related mobility reduction: heterogeneous effects on sleep and physical activity rhythms. *Sleep* 2021 Feb 12;44(2):zsa179 [[FREE Full text](#)] [doi: [10.1093/sleep/zsa179](https://doi.org/10.1093/sleep/zsa179)] [Medline: [32918076](https://pubmed.ncbi.nlm.nih.gov/32918076/)]
66. Pépin JL, Bruno RM, Yang R, Vercamer V, Jouhaud P, Escourrou P, et al. Wearable activity trackers for monitoring adherence to home confinement during the COVID-19 pandemic worldwide: data aggregation and analysis. *J Med Internet Res* 2020 Jun 19;22(6):e19787 [[FREE Full text](#)] [doi: [10.2196/19787](https://doi.org/10.2196/19787)] [Medline: [32501803](https://pubmed.ncbi.nlm.nih.gov/32501803/)]
67. Huckins JF, daSilva AW, Wang W, Hedlund E, Rogers C, Nepal SK, et al. Mental health and behavior of college students during the early phases of the COVID-19 pandemic: longitudinal smartphone and ecological momentary assessment study. *J Med Internet Res* 2020 Jun 17;22(6):e20185 [[FREE Full text](#)] [doi: [10.2196/20185](https://doi.org/10.2196/20185)] [Medline: [32519963](https://pubmed.ncbi.nlm.nih.gov/32519963/)]

## Abbreviations

**MFIS-5:** Modified Fatigue Impact Scale-5

**MS:** multiple sclerosis

**MSRS-R:** Multiple Sclerosis Rating Scale—Revised

**PHQ-9:** Patient Health Questionnaire-9

**PSQI:** Pittsburgh Sleep Quality Index

*Edited by J Torous; submitted 06.04.22; peer-reviewed by N Marotta, N Chiaravalloti; comments to author 27.06.22; revised version received 15.07.22; accepted 16.07.22; published 24.08.22*

*Please cite as:*

Chikersal P, Venkatesh S, Masown K, Walker E, Quraishi D, Dey A, Goel M, Xia Z

*Predicting Multiple Sclerosis Outcomes During the COVID-19 Stay-at-home Period: Observational Study Using Passively Sensed Behaviors and Digital Phenotyping*

*JMIR Ment Health* 2022;9(8):e38495

URL: <https://mental.jmir.org/2022/8/e38495>

doi: [10.2196/38495](https://doi.org/10.2196/38495)

PMID: [35849686](https://pubmed.ncbi.nlm.nih.gov/35849686/)

©Perna Chikersal, Shruthi Venkatesh, Karman Masown, Elizabeth Walker, Danyal Quraishi, Anind Dey, Mayank Goel, Zongqi Xia. Originally published in *JMIR Mental Health* (<https://mental.jmir.org>), 24.08.2022. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in *JMIR Mental Health*, is properly cited. The complete bibliographic information, a link to the original publication on <https://mental.jmir.org/>, as well as this copyright and license information must be included.