Original Paper

Evaluating the Utility of Smartphone-Based Sensor Assessments in Persons With Multiple Sclerosis in the Real-World Using an App (elevateMS): Observational, Prospective Pilot Digital Health Study

Abhishek Pratap[1], PhD; Daniel Grant[2], PhD; Ashok Vegesna[2], PharmD, MS, RPh; Meghasyam Tummalacherla[1], MS; Stanley Cohan[3], MD, PhD; Chinmay Deshpande[2], PhD; Lara Mangravite[1], PhD; Larsson Omberg[1], PhD

[1]Sage Bionetworks, Seattle, WA, United States

[2]Novartis Pharmaceuticals Corporation, East Hanover, NJ, United States

[3]Providence Multiple Sclerosis Center, Providence St Vincent Medical Center, Portland, OR, United States

Corresponding Author:
Larsson Omberg, PhD
Sage Bionetworks
2901 Third Avenue, Suite 330
Seattle, WA, 98121
United States
Phone: 1 (206) 928 8260
Email: larsson.omberg@sagebionetworks.org

Abstract

Background: Multiple sclerosis (MS) is a chronic neurodegenerative disease. Current monitoring practices predominantly rely on brief and infrequent assessments, which may not be representative of the real-world patient experience. Smartphone technology provides an opportunity to assess people's daily-lived experience of MS on a frequent, regular basis outside of episodic clinical evaluations.

Objective: The objectives of this study were to evaluate the feasibility and utility of capturing real-world MS-related health data remotely using a smartphone app, "elevateMS," to investigate the associations between self-reported MS severity and sensor-based active functional tests measurements, and the impact of local weather conditions on disease burden.

Methods: This was a 12-week, observational, digital health study involving 3 cohorts: self-referred participants who reported an MS diagnosis, clinic-referred participants with neurologist-confirmed MS, and participants without MS (controls). Participants downloaded the elevateMS app and completed baseline assessments, including self-reported physical ability (Patient-Determined Disease Steps [PDDS]), as well as longitudinal assessments of quality of life (Quality of Life in Neurological Disorders [Neuro-QoL] Cognitive, Upper Extremity, and Lower Extremity Function) and daily health (MS symptoms, triggers, health, mobility, pain). Participants also completed functional tests (finger-tapping, walk and balance, voice-based Digit Symbol Substitution Test [DSST], and finger-to-nose) as an independent assessment of MS-related cognition and motor activity. Local weather data were collected each time participants completed an active task. Associations between self-reported baseline/longitudinal assessments, functional tests, and weather were evaluated using linear (for cross-sectional data) and mixed-effects (for longitudinal data) regression models.

Results: A total of 660 individuals enrolled in the study; 31 withdrew, 495 had MS (n=359 self-referred, n=136 clinic-referred), and 134 were controls. Participation was highest in clinic-referred versus self-referred participants (median retention: 25.5 vs 7.0 days). The top 5 most common MS symptoms, reported at least once by participants with MS, were fatigue (310/495, 62.6%), weakness (222/495, 44.8%), memory/attention issues (209/495, 42.2%), and difficulty walking (205/495, 41.4%), and the most common triggers were high ambient temperature (259/495, 52.3%), stress (250/495, 50.5%), and late bedtime (221/495, 44.6%). Baseline PDDS was significantly associated with functional test performance in participants with MS (mixed model–based estimate of most significant feature across functional tests [β]: finger-tapping: β=–43.64, P<.001; DSST: β=–5.47, P=.005; walk and balance: β=–.39, P=.001; finger-to-nose: β=.01, P=.01). Longitudinal Neuro-QoL scores were also significantly associated with functional tests (finger-tapping with Upper Extremity Function: β=.40, P<.001; walk and balance with Lower Extremity

XSL•FO
RenderX

Function: β=–99.18, *P*=.02; DSST with Cognitive Function: β=1.60, *P*=.03). Finally, local temperature was significantly associated with participants' test performance (finger-tapping: β=–.14, *P*<.001; DSST: β=–.06, *P*=.009; finger-to-nose: β=–53.88, *P*<.001).

Conclusions: The elevateMS study app captured the real-world experience of MS, characterized some MS symptoms, and assessed the impact of environmental factors on symptom severity. Our study provides further evidence that supports smartphone app use to monitor MS with both active assessments and patient-reported measures of disease burden. App-based tracking may provide unique and timely real-world data for clinicians and patients, resulting in improved disease insights and management.

KEYWORDS

Introduction

Multiple Sclerosis

Multiple sclerosis (MS) is a chronic neurodegenerative disease that affects more than 2 million people worldwide, with prevalence rates equating to greater than 400,000 cases in the United States [1,2]. Symptoms of MS can affect motor function, sensation, cognition, and mood [3], and substantially impact quality of life (QoL) [4]. The combination of MS symptoms, their severity, and the course of disease varies between individual patients and can be affected by environmental factors, such as temperature, vitamin D stores, stress [1,5-10], and comorbidities, including depression, diabetes, cardiovascular disease, cancer, and autoimmune conditions [11,12]. Despite the heterogeneity of MS, certain symptoms and triggers are common among patients; for example, fatigue is reported as a symptom in approximately 75% of patients [13], and elevated temperature is estimated to cause transient symptom worsening in up to 80% [9].

Patient Monitoring and Digital Tools

Routine clinical care for MS typically involves brief assessments performed during infrequent neurologist visits and, hence, often relies on retrospective self-reporting of symptoms and treatment responses, which can be subject to recall bias or affected by MS-associated cognitive impairment [14-17]. As such, this approach may fail to fully capture an individual's day-to-day experience of living with MS and contribute to reduced accuracy and timeliness in detecting changes in symptom burden, disease severity/relapses, therapeutic outcomes, and the need for timely therapeutic agent change [17,18], Furthermore, while patient-reported outcomes (PROs) are increasingly used in MS clinical care, there is a lack of universal guidance on MS-specific PROs, and their usage and interpretation can differ between individual clinicians [17]. These challenges in monitoring MS highlight an unmet need for more effective, patient-centered tools that are able to capture the daily lived experience of disease outside of episodic clinic visits [19].

While a number of studies have used web-based tools to collect MS health data through patient diaries and electronic PROs [20,21], there is a growing need to tailor these digital health tools to the needs of patients with MS; this includes developing sensor-based assessments of MS disease severity and evaluating the impact of environmental factors, such as weather, stress, and sleep impairment on MS burden [22,23]. By bridging the gap between episodic clinical observation and the real-world experience of MS, these digital tools have the potential to improve self-reporting and disease monitoring, and provide a more comprehensive assessment of disease trajectories. In turn, this may support clinicians in making disease management recommendations to patients with MS, as well as other diseases that have heterogeneous and variable symptoms over time [24].

The ubiquity of smartphones with built-in sensors provides an opportunity to address the growing need for real-time disease monitoring. A number of previous studies have demonstrated the feasibility of smartphones in collecting health data in a real-time, real-world setting from patients across a range of disease areas including asthma, diabetes, depression, and Parkinson disease [25-30]. In addition to monitoring symptoms and triggers, these studies have identified geographic and environmental factors related to disease severity, and have been reported by patients to have a positive impact on their disease management [26,27,29]. Several digital health studies have already been undertaken in patients with MS with encouraging results; data suggest that smartphone technology can be effectively leveraged to monitor MS symptom severity, QoL, and medication usage, enabling patients to play an active role in disease management [23,31-35]. The recent FLOODLIGHT study has also shown that smartphone-based active testing can be used to remotely monitor motor function and capture MS symptoms, thus providing a more accurate assessment of MS in the real world [23,34]. However, a number of these studies have involved partially remote designs that include scheduled clinic visits at predetermined time points [32,34], which may limit widespread usage and participation. To this end, additional studies are required to build on these existing data and further assess digital health tools in a large, remote population of patients with MS.
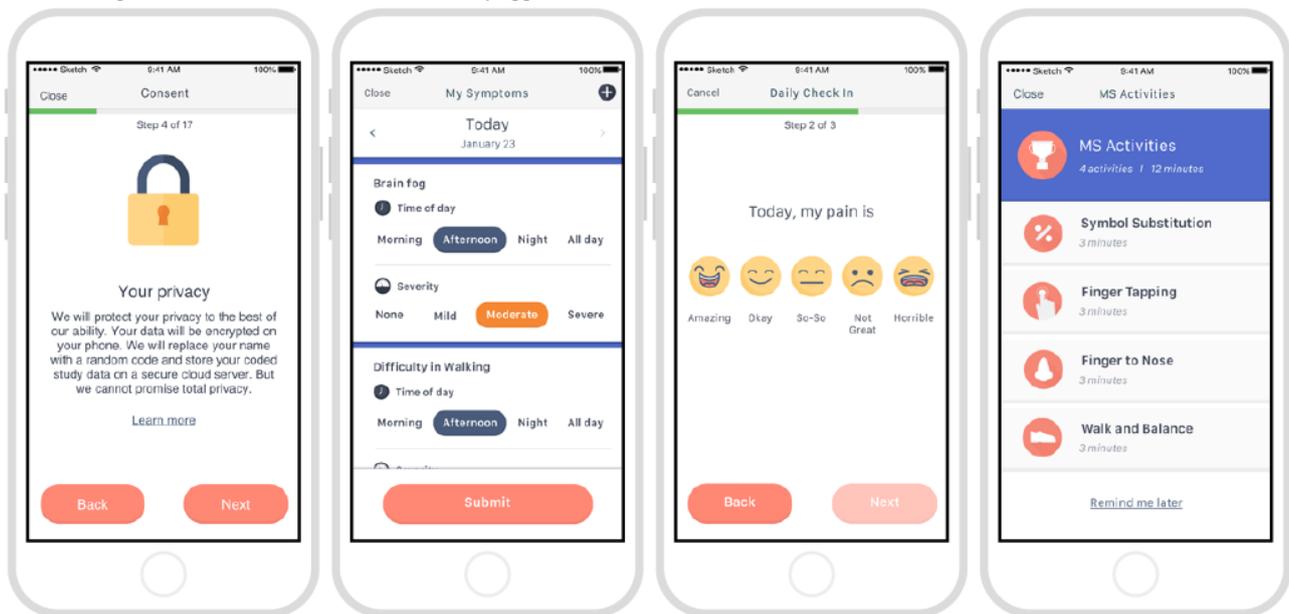
Objective

The main objective of this study was to evaluate the feasibility and utility of gathering MS-related health information from a large, remotely enrolled cohort using a dedicated smartphone app and to monitor study participants over a 12-week period. The study app, "elevateMS," was developed through a user-centered design process. Secondary objectives were to examine the relationship between disease severity and QoL, measured via PROs and performance in sensor-based active functional tests, and to investigate the impact of local weather conditions on variations in MS symptoms and severity.

XSL•FO

RenderX

Methods

Study Design

This was a 12-week observational, prospective pilot digital health study using data collected from a dedicated smartphone app, elevateMS. The app was developed using a patient-centered design process with MS patient advisors (Multimedia Appendix 1) and was freely available to download from the Apple App Store. Enrollment spanned from August 2017 to October 2019, and participants were required to be aged 18 years or older, reside in the United States, and use an iPhone 5 or newer device.

Study Participants

Participants were openly recruited through word of mouth, press releases, online advertisements, and the study website [36] and grouped into 2 cohorts: individuals without MS (controls) and individuals who self-reported diagnosis of MS (self-referred). A third, "clinic-referred" cohort, with a neurologist-confirmed MS diagnosis, was also recruited through information flyers distributed at 3 MS treatment centers. Ethical approval was granted by the Western Institutional Review Board, and enrollment, informed consent, and data collection were carried out electronically through the study app (Figure 1) [37].

Figure 1. Example screenshots from the elevateMS study app.

Participants had to make an active choice to complete the consent process, and no default option was presented. Participants were also given the option to share their data only with the elevateMS study team and partners (share narrowly), or more broadly with qualified researchers worldwide [38].

Data Collection

elevateMS primarily targeted collection of real-world data from participants with MS. This included self-reported measures of symptoms and health via optional "check-in" surveys, and independent assessments of motor function via sensor-based active functional tests. Participants were encouraged to complete check-in surveys on a daily basis and were notified to perform more comprehensive functional tests once a week. Local weather data were collected every time an assessment was performed. The data collected through elevateMS and the frequency at which each element was recorded are summarized in Table 1. With the exception of overall physical ability, which was a baseline-only assessment, all data were collected longitudinally at various intervals over the 12-week study duration.

Table 1. Data collected through the elevateMS study app.

Data source	Timeline of data collection
Baseline demographics	
Sociodemographic data (age, gender, race, education, health insurance, employment status, geographic location)	Day 2
MS[a] disease characteristics (diagnosis, medication, family history)	Day 1
Patient-reported outcomes	
Overall physical ability[b]	Day 2
Check-in survey: MS symptoms and triggers	Daily
Check-in survey: health, mobility, pain[c]	Daily
Short-form Neuro-QoL[d] domains	Every third functional test targeting that domain (Cognitive Function domain: DSST; Upper Extremity Function domain: finger-tapping/finger-to-nose; Lower Extremity Function domain: walk and balance)
Active functional test	
Finger-tapping	Weekly[f]
Walk and balance	Weekly[f]
DSST[e]	Weekly[f]
Finger-to-nose	Weekly[f]
Local weather data	
Temperature	Every time a test was performed
Humidity	Every time a test was performed
Cloud coverage	Every time a test was performed
Atmospheric pressure	Every time a test was performed

[a]MS: multiple sclerosis.

[b]Based on a truncated 4-point Patient-Determined Disease Steps scale, administered at baseline to all patients with MS (Normal, Mild Disability, Moderate Disability, Gait Disability).

[c]Based on a 5-point Likert scale (Health: Amazing, Okay, So-So, Not great, Horrible; Mobility: Excellent, Very good, Good, Not great, Horrible; Pain: None, Mild, Moderate, Severe, Horrible).

[d]Neuro-QoL: Quality of Life in Neurological Disorders.

[e]DSST: Digit Symbol Substitution Test.

[f]With option for participants to complete more frequently.

Participants with MS completed optional daily check-in surveys to record their individual symptoms and triggers, and to assess their health, mobility, and pain using a 5-point Likert scale. Self-reported disease severity was determined using a truncated 4-point Patient-Determined Disease Steps (PDDS) scale to assess overall physical ability (Normal, Mild Disability, Moderate Disability, Gait Disability) [39]. The impact of MS on daily function was also assessed using 3 short-form domains of the previously validated Quality of Life in Neurological Disorders measurement tool (Neuro-QoL; Cognitive Function, Upper Extremity Function, and Lower Extremity Function; Multimedia Appendix 2) [40-42]. Raw Neuro-QoL scores were converted to standardized Theta (T) scores using standard scoring protocols [42], and subsequently classified into discrete, clinically relevant categories (Normal, Mild, Moderate, and Severe) based on past research [43-45]. All participants carried out active functional tests, which used smartphone sensors as a proxy for traditional symptom measurements. These tests included:

1. The finger-tapping test, where participants repeatedly tapped between 2 circles with alternating fingers as fast as possible for 20 seconds, to measure dexterity, speed, and abnormality in movement.

2. The walk and balance test, where participants walked for 20 seconds with their iPhone in their pocket, then stood still for 10 seconds, to assess gait, posture, stability, and balance;

3. The voice-controlled Digit Symbol Substitution Test (DSST) [46], where participants used the microphone to record answers to measure cognitive function.

4. The finger-to-nose test, where participants extended their arm while holding their iPhone and then touched the phone screen to their nose repeatedly, to measure kinetic tremor and dysmetria in each hand.

Raw data collected from sensor-based active functional tests were transformed using the mhealthtools package, an open-source feature engineering pipeline [47]. This process generated features for each functional test related to different aspects of a participant's health state; for example, the finger-tapping test comprised over 40 features, including number of taps, frequency, and location drift. See Multimedia Appendix 3 for specific examples and the elevateMS Feature Definitions webpage for a full list of features [48]. In addition to extracting features, the mhealthtools pipeline also filtered out records lacking data [47].

The data contributed by participants, as well as scheduling of in-app active functional tests, were managed using an open-source platform developed and maintained by Sage Bionetworks [49].

Statistical Analyses

Descriptive statistics were used to summarize and compare the baseline demographics and MS disease–related characteristics in the study cohort.

User-engagement data were collected and analyzed to understand participant demographics, retention, and compliance in the study. Retention analysis was performed under the definition that participants were considered active in the study if they completed at least one survey or sensor-based test in a given week. The total duration in the study was determined by the number of days between the first and last test. Participants' weekly compliance was assessed using a more stringent cut off than weekly retention; a participant was considered minimally compliant if he/she completed at least one out of four weekly sensor-based active functional tests. Overall retention (ie, total duration a participant remained in the study) was examined across the 3 cohorts (controls, self-referred participants with MS, and clinic-referred participants with MS) using Kaplan–Meier plots. A log-rank test was used to compare the retention difference between the 3 cohorts. The impact of baseline demographics and MS disease characteristics on participant retention was assessed using a Cox proportional hazards model. Each covariate of interest was tested independently, including an interaction term for the clinical referral status. The assumption of the proportional hazard model was tested using scaled Schoenfeld residuals. Finally, per the study protocol, participants were expected to remain in the study for 12 weeks, and thus all user-engagement analysis was limited to the first 12 weeks of study participation.

Linear regression models were used to test for association between participants' self-reported demographics, baseline physical ability (collected once during the onboarding), and the median value of all features generated for each of the 4 sensor-based active functional tests. To test for association between longitudinal PRO assessments (ie, Neuro-QoL results and daily check-ins) and active functional test performance, as well as the potential impact of local weather conditions, a linear mixed-effects (LME) modeling approach was used to account for the subject-level heterogeneity. Prior to modeling, the PRO data were aligned with more frequently administered functional tests by aggregating the average value for all features per week. LME models were fit using the R package lme4, version 1.1-23

[50] with combinations of fixed and random effects. Because of a significant amount of missing responses in sociodemographic information for participants (see Multimedia Appendix 4), some LME models did not converge, and therefore a simpler LME model accounting for participant-level random effect only was used. Statistical significance (*P*-values) for LME models was determined using the Satterthwaite degrees of freedom method through the lmerTest package (version 3.1-1) [51]. For analysis conducted using LME models, we report estimates (β) of all fixed effects covariates along with *P*-values. In addition, *P*-values from ANOVA test conducted to assess significant differences in LME model fixed effect estimates (regardless of the individual factor levels) are also reported. In case a functional test was due to be completed by both hands, the LME models also accounted for variations due to left- and right-hand differences. As participants with MS completed active functional tests at different frequencies, participation rates, sample size, and number of data points varied between analyses. We also conducted sensitivity analyses to evaluate the impact of extreme Neuro-QoL and health, mobility, and pain categories on association results; this involved excluding functional test scores that mapped to the Severe Neuro-QoL category and excluding health, mobility, and pain scores that mapped to the Horrible category. All *P*-values were corrected for multiple testing and false positives using the Benjamini–Hochberg procedure. All analyses were performed using open-source statistical analysis framework R (version 3.5.2; R Foundation for Statistical Computing) [52].

Data Availability

Complete results from this analysis are available online through the accompanying elevateMS study portal [53]. Additionally, individual user-level raw data for those participants who consented to share their data broadly with qualified researchers worldwide is also available under controlled access through the study portal [54].

Results

Study Population

The elevateMS app was released in August 2017 through the Apple App Store and enrolled participants on a rolling basis until October 2019. A total of 660 participants enrolled in the study, of which 31 selected to withdraw with no reason provided. Of the remaining 629 participants, 134 (21.3%) were controls (self-reported as not having MS) and 495 (78.7%) were participants with MS. Of the 495 participants with MS, 359 (72.5%) self-referred to the study with a self-reported MS diagnosis and 136 (27.5%) were referred from 3 clinical sites and had a neurologist-confirmed MS diagnosis. Participants were located across the United States (Figure 2), with a mean (SD) age of 39.34 (11.41), 45.20 (11.64), and 48.93 (11.20) years in the control, self-referred, and clinic-referred cohorts, respectively. A summary of baseline sociodemographic data is presented in Table 2. Further information on missing responses in demographic data is presented in Multimedia Appendix 4.

Baseline disease characteristics for the self-referred and clinic-referred participants with MS are shown in Table 3. Most participants reported relapsing–remitting MS; this included

83.6% (300/359) of the self-referred cohort and 90.4% (123/136) of the clinic-referred cohort. Infusion disease-modifying therapy was the most common treatment received by both self-referred (116/359, 32.3%) and clinic-referred (64/136, 47.1%) participants.

Figure 2. Geographic locations of participants. Dots (n=329) represent the location of those participants who continued in the study beyond initial enrollment and provided the first three digits of their zip code during the collection of demographic information on Day 2. One dot is included for each location, with participants with the same first three digits of the zip code shown under the same dot.
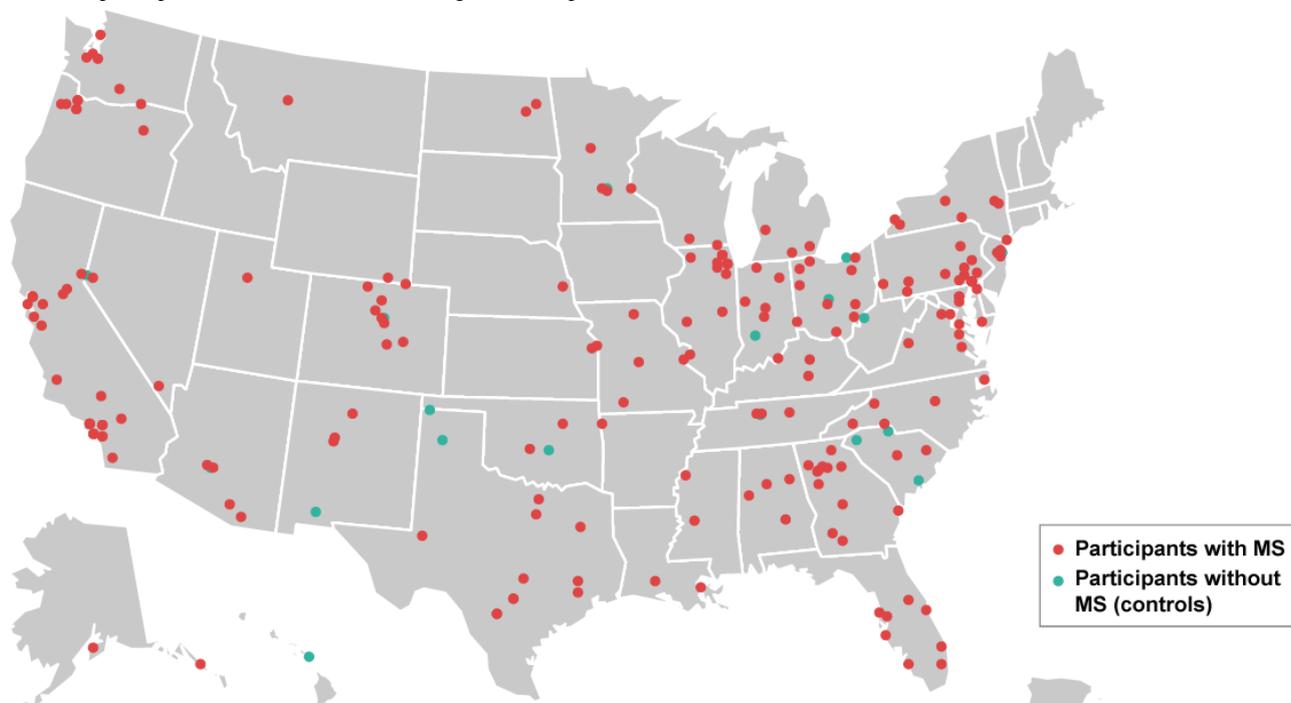
Table 2. Baseline sociodemographic characteristics of study participants.[a]

Characteristic	Controls (N=134)	Participants with MS[b] (self-referred; N=359)	Participants with MS (clinic-referred; N=136)
Age (years), mean (SD)	39.34 (11.41)	45.20 (11.64)	48.93 (11.20)
Gender			
Female	27 (64.3)	154 (73.3)	78 (84.8)
Male	15 (35.7)	56 (26.7)	14 (15.2)
Race			
Asian	6 (13.6)	4 (1.9)	0 (0.0)
Black African	1 (2.3)	13 (6.1)	9 (9.8)
Caucasian	26 (59.1)	182 (85.4)	74 (80.4)
Latino Hispanic	5 (11.4)	9 (4.2)	5 (5.4)
Other	6 (13.6)	5 (2.3)	4 (4.3)
Education			
College degree	16 (37.2)	123 (57.5)	55 (60.4)
High-school diploma/GED[c]	4 (9.3)	16 (7.5)	7 (7.7)
Postgraduate degree	23 (53.5)	70 (32.7)	29 (31.9)
Other	0 (0.0)	5 (2.3)	0 (0.0)
Health insurance			
Government insurance	3 (7.0)	65 (30.5)	19 (20.7)
Employer insurance	30 (69.8)	100 (46.9)	55 (59.8)
No insurance	1 (2.3)	2 (0.9)	2 (2.2)
Other	9 (20.9)	46 (21.6)	16 (17.4)
Employment status			
Full-time	30 (69.8)	93 (43.5)	41 (44.6)
Part-time	3 (7.0)	17 (7.9)	11 (12.0)
Retired	3 (7.0)	18 (8.4)	10 (10.9)
Disabled	4 (9.3)	62 (29.0)	19 (20.7)
Unemployed	0 (0.0)	10 (4.7)	3 (3.3)
Other	3 (7.0)	14 (6.5)	8 (8.7)

[a]All data shown are n (%), unless otherwise stated. Percentages were calculated based on the total number of participants who provided a response and excluded missing information. See Multimedia Appendix 4 for further details on missing results, including the number and proportion of participants who did not provide responses.

[b]MS: multiple sclerosis.

[c]GED: General Educational Development.

Table 3. Baseline disease characteristics of study participants with MS.

Characteristic[a]	Participants with MS (self-referred; N=359)	Participants with MS (clinic referred; N=136)
MS[b] diagnosis		
Relapsing–remitting	300 (83.6)	123 (90.4)
Primary progressive	34 (9.5)	6 (4.4)
Secondary progressive	25 (7.0)	5 (3.7)
Not sure	0 (0.0)	2 (1.5)
Current DMT[c]		
Infusion	116 (32.3)	64 (47.1)
Injection	83 (23.1)	24 (17.6)
Oral	114 (31.8)	40 (29.4)
None	46 (12.8)	6 (4.4)
Missing	0 (0.0)	2 (1.5)
MS family history		
Yes	77 (21.4)	23 (16.9)
No	251 (69.9)	104 (76.5)
Not sure	31 (8.6)	9 (6.6)
Overall physical ability[d]		
Normal	101 (28.1)	52 (38.2)
Gait disability	85 (23.7)	24 (17.6)
Mild disability	104 (29.0)	38 (27.9)
Moderate disability	69 (19.2)	20 (14.7)
Missing	0 (0.0)	2 (1.5)
Duration of disease		
Years since diagnosis, mean (SD)	11.14 (8.86)	14.29 (8.89)
Duration of treatment		
Years since first DMT, mean (SD)	10.09 (7.97)	13.07 (7.92)

[a]All data shown are n (%), unless otherwise stated.

[b]MS: multiple sclerosis.

[c]DMT: disease-modifying therapy.

[d]Based on truncated 4-point Patient Determined Disease Steps scale.

Participant Engagement

Median study retention was significantly higher in clinic-referred participants with MS (25.5 days [95% CI 17.0-55.0]) compared with self-referred participants with MS (7.0 days [95% CI 4.0-11.0]) and controls (1.0 day [95% CI 1.0-2.0 days]; $P<.001$; Figure 3). Compliance, defined in this study as the completion of at least one sensor-based active functional test per week, decreased over time in all cohorts; from Week 1 to Week 12,

compliance fell from 80.2% to 50.0% for the clinic-referred cohort, from 81.1% at Week 1 to 46.1% for the self-referred cohort, and from 70.9% to 20.0% in the control cohort (Figure 3). Given the lack of ongoing engagement in the control group and the fact that elevateMS primarily targeted participants with MS, the control cohort was not included in subsequent data analyses, and results from self-referred and clinic-referred participants with MS were pooled for analysis.

Figure 3. elevateMS user engagement. Participant retention (median number of days in the study) and compliance (completion of at least one out of four sensor-based active functional tests per week) across the three study cohorts. MS: multiple sclerosis.
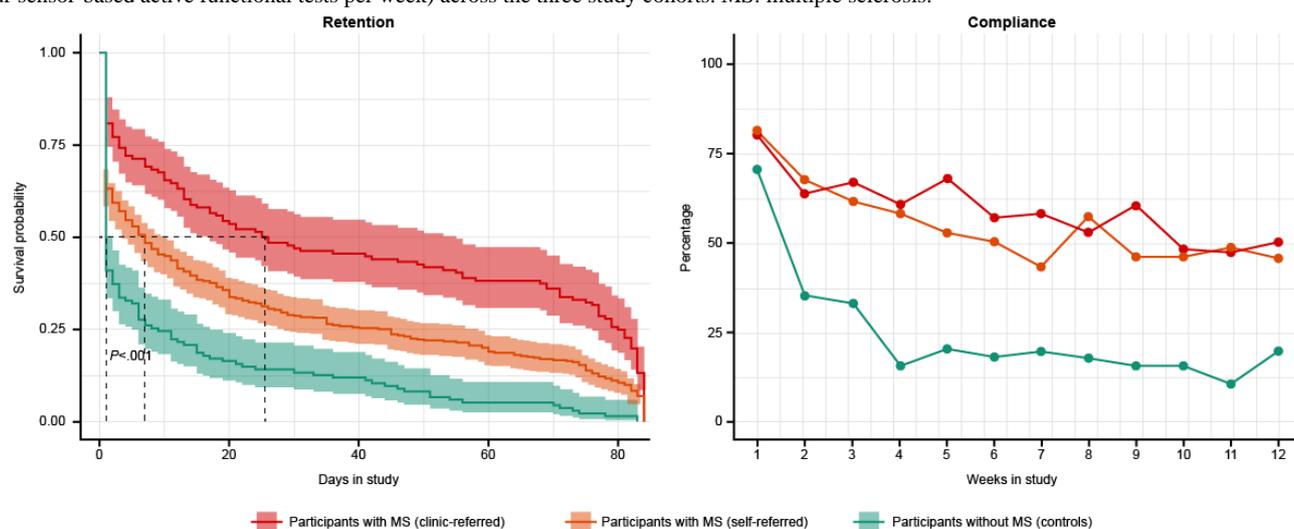

For patients with MS, participation in all tasks (both sensor-based active functional tests and daily check-in surveys) was highest in Weeks 1 and 2, then decreased over time (Multimedia Appendix 5). When using the elevateMS app, participants with MS completed the active functional tests (median 40% of overall individual activity; IQR 30.3) the most, followed by reporting MS symptoms and triggers (median 33.3% of overall individual activity; IQR 33) and completing daily check-in surveys (median 22.3% of overall individual activity; IQR 18.2; Multimedia Appendix 6). The most common self-reported symptoms were fatigue, weakness, memory/attention issues, and difficulty walking, and the most common self-reported triggers were high ambient temperature, emotional stress, and going to bed late, all of which were experienced on at least one occasion by more than half of participants with MS (Multimedia Appendix 7). Data collected from sensor-based active tests were transformed into features and filtered for validity using the mhealthtools pipeline [47] (see Multimedia Appendix 8 for further details).

Relationship Between Baseline Physical Ability and Performance in Active Functional Tests

Higher physical disability at baseline was associated with significantly worse performance ($P<.001$; top associations listed below) in multiple sensor-based active functional tests. For each active functional test, a subset of the most significantly associated features is presented here (Figure 4 and Multimedia Appendix 9); the full list of per-feature results is available via Synapse, an online data repository managed by Sage Bionetworks [55]. For the finger-tapping assay, a reduced number of finger taps was statistically significantly associated with baseline physical ability category ($\beta_{\text{gait disability vs normal}}=-43.64$, $P<.001$). Similarly, fewer correct responses in the voice-based DSST task was significantly associated with baseline physical ability ($\beta_{\text{gait disability vs normal}}=-5.47$, $P=.005$). Worse performance in the walking test, based on a feature derived from the device accelerometer (F0FAJ) [48], was statistically significantly associated with low baseline physical ability ($\beta_{\text{gait disability vs normal}}=-0.39$, $P=.001$). For the finger-to-nose test, a tremor feature derived from capturing the hand rotation velocity was also found to be significantly associated with baseline physical ability ($\beta_{\text{gait disability vs normal}}=0.01$, $P=.01$). Balance features were not significantly associated with baseline physical ability (data not shown). Notably, the feature most associated with performance from the finger-tapping assay (median number of finger taps) was also significantly associated with several sociodemographic characteristics in participants with MS, including age group ($P<.001$), education ($P=.001$), duration of treatment ($P=.004$), and duration of disease ($P=.009$; Multimedia Appendix 9). Additionally, baseline physical ability was significantly associated with duration of treatment ($P=.003$), duration of disease, employment status, type of health insurance, and age group (all $P<.001$) in participants with MS (Multimedia Appendix 10).

Figure 4. Association between baseline characteristics and functional test performance in participants with MS. DSST (Digit Symbol Substitution Test): decrease in number of correct DSST responses with increased baseline physical disability; F0FAJ: frequency at which the maximum peak of the Lomb-Scargle periodogram occurred for the average acceleration series, with frequencies limited to 0.2-5 Hz; Finger-tapping: decrease in median number of finger taps with increased baseline physical disability; Finger-to-nose: increase in hand rotation velocity tremor feature with increased baseline physical disability; MS: multiple sclerosis; PDDS: Patient-Determined Disease Steps; Walking: decrease in F0FAJ accelerometer results with increased baseline physical disability.
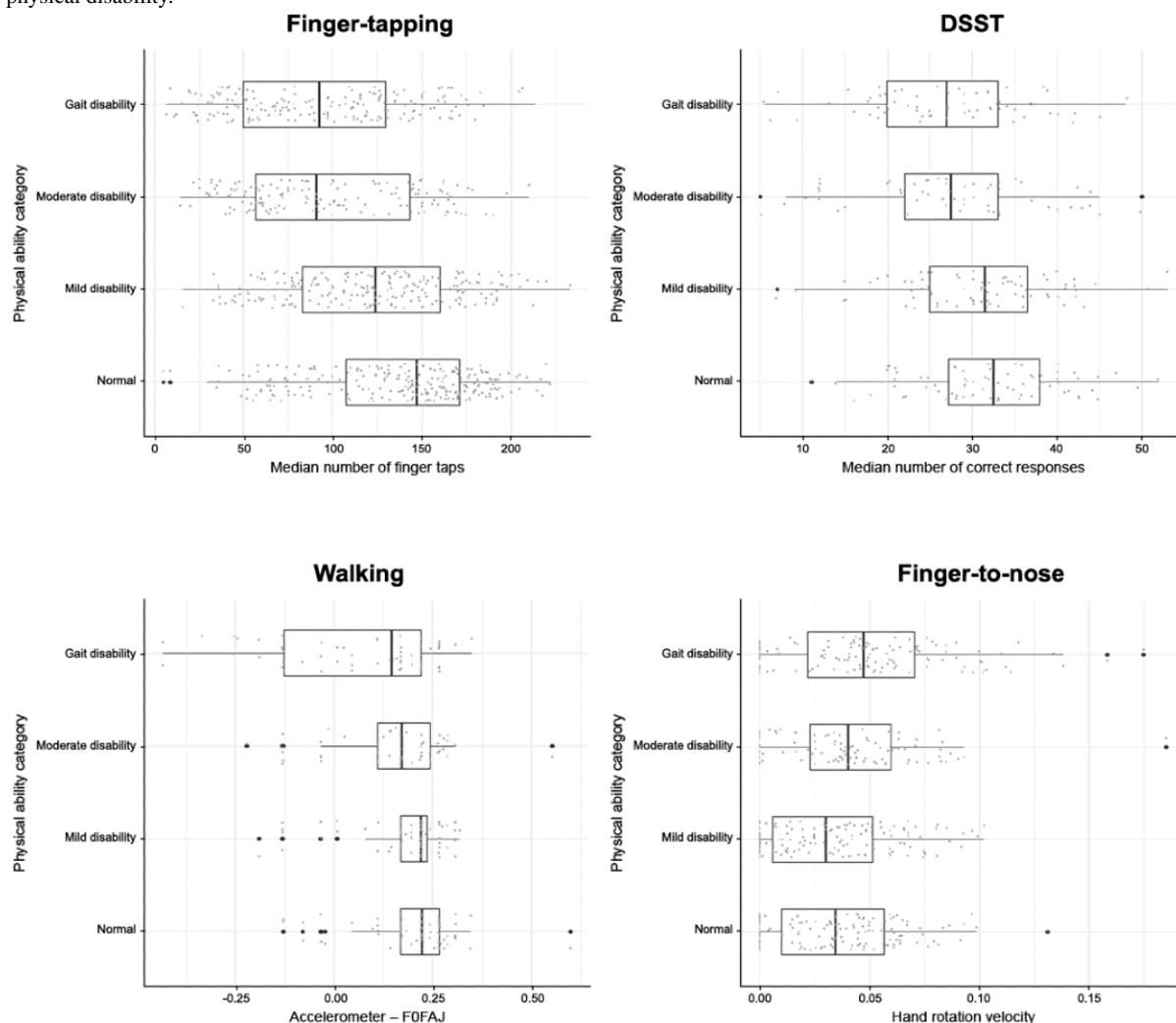
Relationship Between Neuro-QoL Domains and Performance in Active Functional Tests

Performance in sensor-based active functional tests was also significantly associated ($P<.001$; top associations listed below) with the 3 short-form Neuro-QoL surveys (Cognitive Function, Upper Extremity Function, and Lower Extremity Function) that were administered to participants at several points during the course of the study. For each test, a subset of the top most significantly associated features are presented here (Figure 5 and Multimedia Appendix 11), with the full list of per-feature comparisons available online via Synapse data repository [56]. Performance in the finger-tapping test was significantly associated with Upper Extremity Function domain scores ($\beta_{moderate\ vs\ normal}$=0.40 seconds, $P<.001$) of the Neuro-QoL. For the walking test, a feature derived from the device accelerometer (P0Y) [48] was found to be significantly associated with the Neuro-QoL Lower Extremity Function domain scores ($\beta_{moderate\ vs\ normal}$=–99.18 seconds, P=.02). Increased severity in the

Neuro-QoL Cognitive and Lower Extremity Function domains was also significantly associated with the variation in time taken to complete the voice-based DSST task (Cognitive Function: $\beta_{severe\ vs\ normal}$=1.60 seconds, P=.03; Lower Extremity Function: $\beta_{severe\ vs\ normal}$=10.31 seconds, $P<.001$). Finally, for the finger-to-nose test, a feature derived from capturing the hand rotation acceleration using the device gyroscope was significantly associated with the Neuro-QoL Upper Extremity Function domain ($\beta_{moderate\ vs\ normal}$=.11, P=.003). Additional sensitivity analyses were performed to assess the impact of Neuro-QoL scores that mapped to the Severe Neuro-QoL category. While excluding Neuro-QoL scores in severe category did not change the main findings for the finger-tapping, walk and balance, and finger-to-nose tests, the analysis showed that the significant association of Cognitive Function and Lower Extremity Function with DSST was mainly driven by those participants reporting severe Neuro-QoL outcomes (Multimedia Appendix 11).

Figure 5. Association between Neuro-QoL^{TM} domains and functional test performance in participants with MS. Neuro-QoL categories comprising <5% of total participants were not plotted. DSST (Digit Symbol Substitution Test): increase in DSST response time with increased severity in Neuro-QoL Cognition domain. Finger-tapping: increase in maximum finger tapping interval with increased severity in the Neuro-QoL Upper Extremity Function domain; Finger-to-nose: increase in hand rotation acceleration tremor feature with increased severity in Neuro-QoL Upper Extremity Function domain; MS: multiple sclerosis; Neuro-QoL: Quality of Life in Neurological Disorders; P0Y: maximum power in the inspected frequency interval of the Lomb–Scargle periodogram for the Y acceleration series (0.2–5 Hz); Walking: decrease in accelerometer-derived feature (P0Y) with increased severity in Neuro-QoL Lower Extremity Function domain.
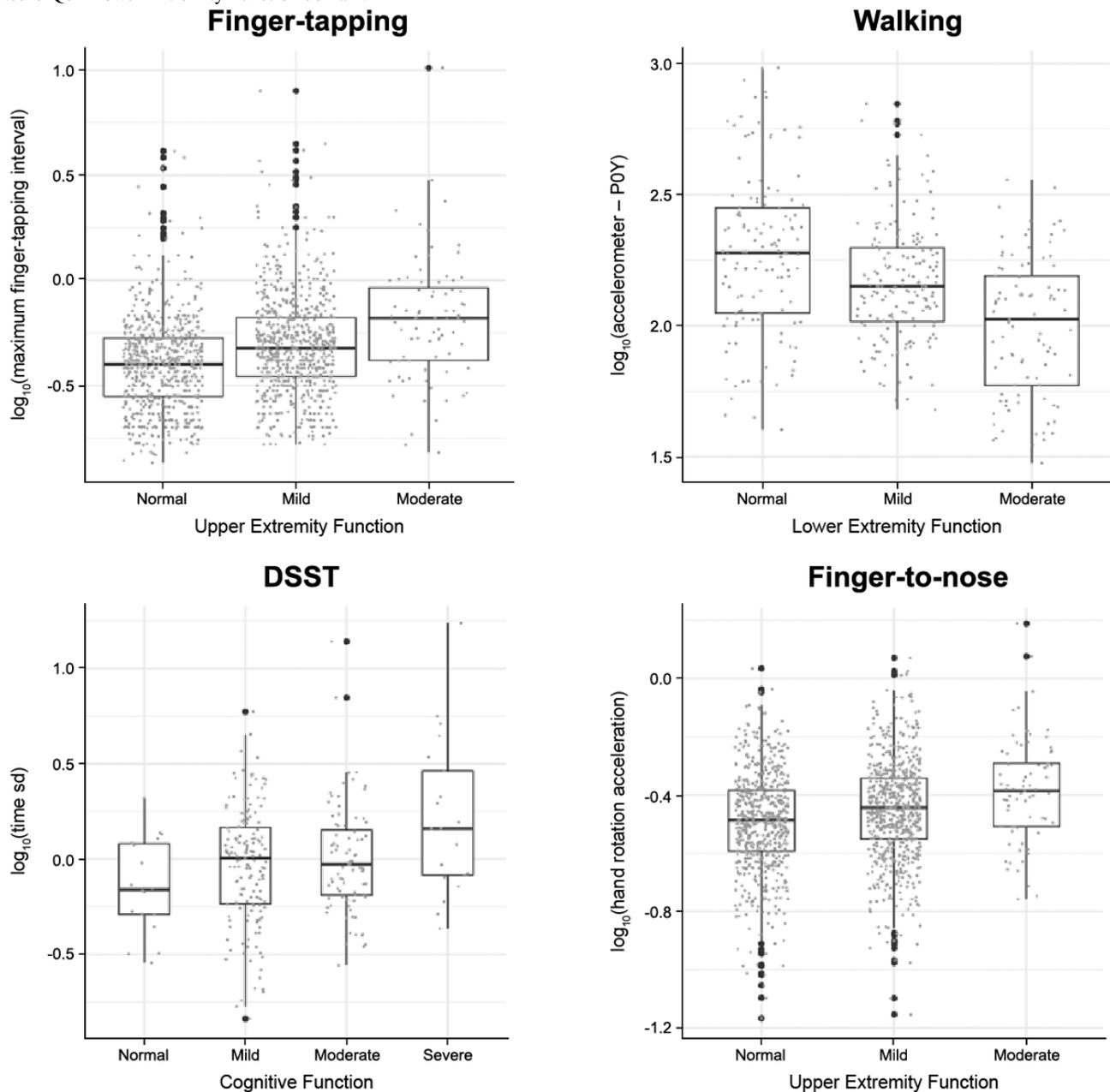
Impact of Local Weather Conditions on Performance in Active Functional Tests and Daily Check-In PROs

Local weather data were captured each time a participant completed a functional test and were found to be significantly associated (*P*<.001; top associations listed below) with respective test performance (Figure 6 and Multimedia Appendix 12) [57]. The local temperature elevations at the time of test completion negatively impacted participant's performance to the greatest extent, with a significant difference in performance in the finger-tapping test observed (β=–.14, *P*<.001). For example, with a 30°F increase in temperature (ie, from 50°F to

80°F), the participant's performance dropped by an average of 4.3 finger-taps (β × 30°F). Similarly, performance in the voice-based DSST and finger-to-nose tests was significantly associated with increased temperature (β=–.06, *P*=.009 and β=–53.88, *P*<.001, respectively). In contrast to active functional test features, PROs recorded from daily check-in surveys were only moderately associated with local weather conditions (Multimedia Appendix 13) [58]. Sensitivity analysis further showed that the association between PROs and weather features was mainly driven by a small proportion of participants reporting extreme outcomes on the Likert scale (Multimedia Appendix 14) [59].

Figure 6. Association between daily temperature and functional test performance in participants with MS. Finger-tapping: decrease in number of finger taps with increased temperature. DSST (Digit Symbol Substitution Test): decrease in number of correct DSST responses with increased temperature. Finger-to-nose: decrease in accelerometer-derived linear jerk tremor feature derived with increased temperature. MS: multiple sclerosis.
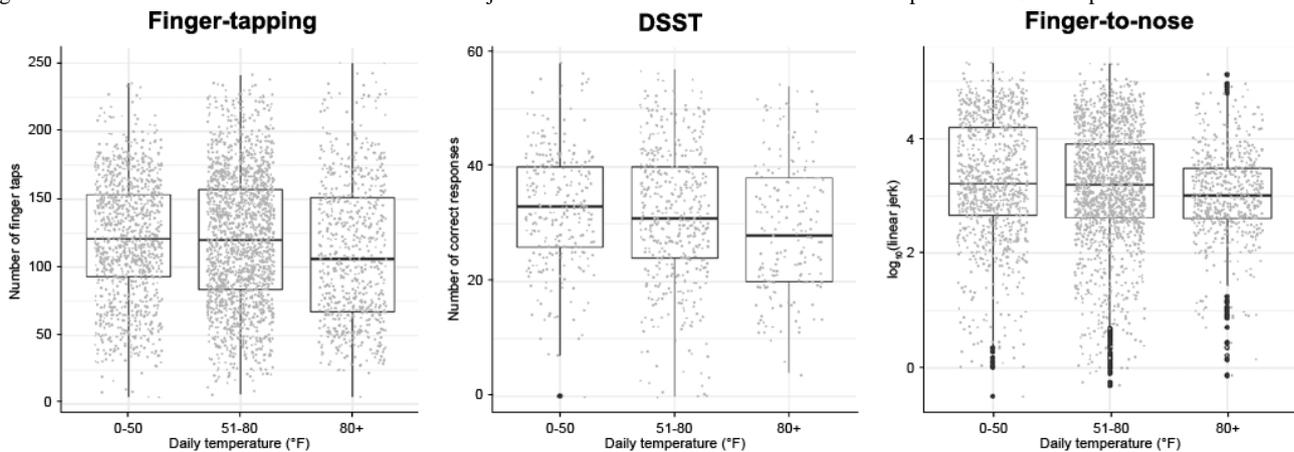
Discussion

Principal Results

The results from this observational, remote data collection study demonstrate the feasibility and utility of a decentralized method to gather real-world data about participant's real-time life experience of MS through a digital health app, elevateMS. Compared with previous digital health studies in MS, elevateMS enrolled one of the largest remote cohorts with a self-reported or neurologist-confirmed diagnosis of MS from across the United States [23,31,32,34,35]. The sociodemographic characteristics of the enrolled patient cohort were broadly similar to the wider population of patients with MS [60,61]. Participation in elevateMS was also generally consistent with that of digital health studies in other disease areas [62]. However, to our knowledge, elevateMS is one of the first remote, digital studies to show the significant impact of clinical referral on overall engagement in a remote population [22]; clinic-referred participants with MS remained active in our study almost 3 times longer than self-referred participants with MS. Importantly, the digitally measured functional activity correlated with clinical outcomes and QoL.

Through the longitudinal collection of PROs, active functional test results, and local weather data, the elevateMS study demonstrates the importance of frequent, real-world assessments of MS disease manifestations outside of episodic clinical evaluations. Tracking of self-reported data identified the most common disease symptoms and triggers in patients with MS, as well as significant associations ($P<.001$) between performance in active functional tests and disease severity, measured by both PDDS and Neuro-QoL subdomain scores. Although PROs failed to capture the impact of local weather conditions, participants' performance in various active functional test features was found to be significantly associated ($P<.001$) with local temperature patterns, with the worst performance observed at temperatures over 80°F. This supports the well-established link between increased temperature and MS [7-9] and demonstrates the sensitivity of these sensor-based tests in patient monitoring. Together, these results show the potential utility of active functional tests in capturing measurements of MS-related motor activity and assessing the impact of local environmental factors on disease symptoms and severity in a real-world setting.

A major strength of the elevateMS study is that it collected self-reported and self-administered measurements of MS health data from patients remotely. PROs are increasingly recognized as valid and meaningful clinical measures for disease monitoring and patient care across a range of therapeutic areas [63-72]. Within the MS field, improved self-assessment and QoL reporting, particularly using electronic/digital tools such as elevateMS in a remote, unsupervised setting, has the potential to benefit both patients and clinicians by enhancing communication and understanding of individual patient needs, thus improving the overall patient experience and informing therapy selection [17,24,73,74]. This is particularly important given that a low level of concordance has been observed between patients and neurologists in recognizing MS relapses, assessing health status, and identifying QoL parameters that are of greatest concern to the individual patient [17,24,75].

Another strength of this study is that both PROs and active measurements were collected contemporaneously in a real-time and real-world setting, with a frequency far exceeding that obtainable in routine MS clinical care. This differentiates elevateMS from the episodic and retrospective periods of data collection that are characteristic of traditional, clinic-based care and often subject to recall bias. By capturing a more comprehensive body of data, such as the frequency of triggers, variability of symptoms, and effect of environmental conditions, elevateMS could enable the interaction between daily life stressors and MS severity to be better evaluated. Furthermore, by leveraging frequent and low-burden assessments, elevateMS and other digital health tools may facilitate regular patient monitoring between clinic visits. This could complement the existing clinical practice, by helping patients to record their symptoms, relapses, and medication usage more accurately and have an active role in their disease management; in turn, this has the potential to provide a more thorough assessment of disease, improve communication with health care providers, and ultimately support clinicians in developing personalized treatment plans [19,27,29,32,34,35]. More broadly and universally applied, this technology may also provide novel insights into the course of chronic and progressive conditions,

as demonstrated by previous digital health studies in MS, asthma, and Parkinson disease [25-27,34]. Finally, as shown by elevateMS, digital health tools can be utilized to gather data from large, remote populations and could, therefore, offer unique opportunities to track and evaluate drug efficacy in a continuous, real-world setting through decentralized clinical trials [34].

Limitations

Some potential limitations need to be considered when interpreting the results from this study. Given that participation required a specific smartphone, the study population may be subject to selection bias [62]; however, the sociodemographic characteristics available from the enrolled patient cohort are broadly representative of the general population with MS [60,61]. As this study largely focused on patients with MS, participation by the control cohort was low, and these data were not included in our analyses. Approximately two-thirds of the total MS cohort self-reported their disease status, and these data may be vulnerable to inaccuracies. However, prior studies have shown that self-reported MS diagnoses in the Pacific Northwest MS Registry were subsequently confirmed by neurology health care providers in more than 98% of cases [76]. In a second study, levels of disability were accurately self-reported by patients and comparable to neurologist ratings in 73% of cases [77]. Within the elevateMS study, data from both self-referred and clinic-referred patients with MS were pooled for analysis, thus precluding between-group comparisons. Future studies could evaluate these 2 patient cohorts separately to enable results to be analyzed in relation to clinical referral status.

We also observed a large degree of missing baseline demographic data within this pilot study, which may reflect the fact that this information was not collected until Day 2 of participation, in an attempt to reduce initial onboarding participant burden. This missing data could impact statistical inference, in particular the robustness of the LME models, as the random effect of many sociodemographic characteristics could not be accounted for. Given that a significant proportion of participants in remote app-based studies drop out during the first week, with the majority leaving on Day 1-2 [62], future studies should prioritize immediate collection of baseline demographic data and make this a compulsory step upon enrollment. This would ensure that demographic information can be fully evaluated in relation to app study results.

Given the limited, 12-week study participation period, we were not able to assess longer-term variations in MS symptoms or severity in this pilot study. Furthermore, this short window of observation meant that disease severity could be assessed in relation only to daily, and not seasonal, fluctuations. Going forward, longer study durations would provide greater opportunities for longitudinal disease monitoring and assessment of the impact of external lifestyle and environmental factors. This is particularly important in chronic diseases such as MS, where there is a risk of progression and unpredictable variability of symptoms or relapses over time [5].

The sample size and engagement of the elevateMS patient cohort were generally consistent with or higher than those in previously published remote digital studies undertaken across different therapy areas [28-30,62]. However, the overall participation in elevateMS was still low compared with digital studies in MS that included an in-person clinic visit [31,32]. The greatest proportion of activity within the elevateMS study app was spent on sensor-based active functional tests, followed by symptom and trigger surveys, and check-in surveys. Participation in specific app activities may have been affected by the time, effort, or frequency associated with that activity; for example, each of the 4 sensor-based active tests were administered on a weekly basis and could be completed in less than 1 minute, whereas each survey was administered daily, potentially making them more burdensome despite their short length. Future studies could further align the study protocol with participant needs and busy schedules employing user-centered co-design techniques. Review of compliance patterns also demonstrated that patients referred from and currently under the care of a provider were the most adherent participants, and future studies should focus on building studies around this cohort. Doing so will not only increase the likelihood of patient retention, but will also provide physician verification of the validity of responses. Additional studies are necessary to develop strategies, including the use of incentives to increase involvement and adherence of people who self-identify and self-refer, to expand the number of patient participants with greater assurance of long-term adherence.

Within the elevateMS study, clinic-referred participants demonstrated greater retention and compliance than the self-referred cohort, which suggests that participants are more likely to engage if they are encouraged by a clinician or aware of how their data may be used to inform and personalize their care. Based on these results, future digital health studies should be incorporated, as exploratory outcomes initially, within clinical trials to assess the applicability and utility of digital monitoring. Finally, it is known that comorbidities, as well as concurrent medication use, have a significant impact on MS patients [11,12,78,79]; details regarding comorbidities were not collected in this pilot study, but should be included in future long-term assessments.

Comparison With Prior Work

In line with previous digital health feasibility studies, our results demonstrate that smartphone technology can be used to collect both sensor-based active measurements and passive data related to disease symptoms and severity in patients with MS [23,31-35]. In contrast to previous studies, elevateMS collected data from a large, geographically diverse, remote, and unsupervised population, independent of scheduled clinic visits. Although a significant number of individuals did not participate beyond enrollment on the first day of the elevateMS study, our user engagement data are consistent with previous digital health studies that recruited broadly from the general population, with no scheduled in-clinic touchpoints or incentives associated with the app usage [62]. As with previous digital health studies [26,28,34], elevateMS relied on arbitrary measures of retention and compliance. In order to fully assess participant engagement and enable comparisons between different studies, these parameters need to be defined in more specific terms. For example, the BEST (Biomarkers, EndpointS, and other Tools) Resource, created by the US Food and Drug Administration and the National Institutes of Health [80], could be expanded to include clear and unambiguous definitions of retention and

compliance in digital health studies, creating standardized measures that could be utilized across the field.

Clinical validation of elevateMS data was beyond the scope of this study; however, it is reassuring that the symptoms and triggers most commonly reported in the app, such as fatigue, weakness, temperature and stress, are already well-documented in traditional studies of MS [7-10,81]. This is further reinforced by our results showing worse performance in active functional tests, such as finger-tapping and DSST responses, in increased ambient temperatures, reflecting the well-known heat sensitivity experienced by patients with MS [7-9]. Furthermore, by comparing self-reported PDDS and Neuro-QoL results, we have shown that it is possible to use smartphone-based motor measurements to assess both disease severity and QoL, providing internal validation of elevateMS results.

Conclusions

In contrast to current, episodic disease monitoring practices, this study demonstrates the value and utility of frequently assessing the real-world, live patient experience of MS using a digital health app. By providing a more comprehensive and representative assessment of patients outside of the clinic, elevateMS and other disease-tracking apps have the potential to enhance the understanding of MS, facilitate patient–clinician communication, and support personalization of disease management plans.

Acknowledgments

Authors' Contributions

LO, DG, AV, and CD were involved in the conception, study design, and execution of the investigation. LM acted as investigator and reviewed the results and manuscript. AP oversaw the data curation and featurization and conducted the analysis. LO wrote the initial analytical plan and oversaw the study. MT was responsible for feature extraction from raw sensor-level data as well as the analysis and curation of study data for public release. SC provided expert knowledge to contextualize results from a clinically meaningful perspective and reviewed the results and manuscript. AP and LO co-wrote the first draft of the manuscript and contributed substantially to the final draft.

Conflicts of Interest

SC received consulting fees for research/speaking honoraria/advisory boards from AbbVie, Alexion, Atara Biotherapeutics, Biogen, EMD Serono, Novartis, Roche/Genentech, Sanofi Genzyme, MedDay, and Pear Therapeutics. AP, LM, MT, and LO are employees of Sage Bionetworks. AV and CD are employees of Novartis Pharmaceuticals Corporation; DG is employed by Alcon.

Multimedia Appendix 1

Overview of user-centered design process used to create the elevateMS study app.
[DOCX File , 60 KB-Multimedia Appendix 1]

Multimedia Appendix 2

Example elevateMS survey questions.
[DOCX File , 39 KB-Multimedia Appendix 2]

Multimedia Appendix 3

Example features from active functional performance tests.
[DOCX File , 27 KB-Multimedia Appendix 3]

Multimedia Appendix 4

Missing responses in baseline sociodemographic characteristics of study participants.
[DOCX File , 27 KB-Multimedia Appendix 4]

XSL•FO
RenderX

Multimedia Appendix 5

Summary of activity-specific compliance across the 12-week study duration.
[DOCX File , 39 KB-Multimedia Appendix 5]

Multimedia Appendix 6

Distribution of overall activity type per user for participants with multiple sclerosis.
[PNG File , 36 KB-Multimedia Appendix 6]

Multimedia Appendix 7

Participant-reported symptoms and triggers.
[DOCX File , 35 KB-Multimedia Appendix 7]

Multimedia Appendix 8

Example assessment of participant performance in the walk and balance sensor-based active functional test.
[DOCX File , 146 KB-Multimedia Appendix 8]

Multimedia Appendix 9

Association between baseline characteristics and active functional test performance in participants with MS (top features for each test).
[DOCX File , 45 KB-Multimedia Appendix 9]

Multimedia Appendix 10

Heatmap showing association between all recorded baseline characteristics in participants with multiple sclerosis.
[DOCX File , 61 KB-Multimedia Appendix 10]

Multimedia Appendix 11

Association between Neuro-QoL™ domains and functional test performance in participants with MS (top features for each test).
[DOCX File , 48 KB-Multimedia Appendix 11]

Multimedia Appendix 12

Association between local weather conditions and functional test performance in participants with MS.
[DOCX File , 39 KB-Multimedia Appendix 12]

Multimedia Appendix 13

Association between local weather conditions and PROs in participants with MS.
[DOCX File , 61 KB-Multimedia Appendix 13]

Multimedia Appendix 14

Association between local weather conditions and PROs in participants with MS (sensitivity analysis).
[DOCX File , 58 KB-Multimedia Appendix 14]

References

1. GBD 2015 Neurological Disorders Collaborator Group. Global, regional, and national burden of neurological disorders during 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Neurol 2017 Nov;16(11):877-897 [FREE Full text] [doi: 10.1016/S1474-4422(17)30299-5] [Medline: 28931491]
2. Zwibel HL, Smrtka J. Improving quality of life in multiple sclerosis: an unmet need. Am J Manag Care 2011 May;17 Suppl 5:S139-S145 [FREE Full text] [Medline: 21761952]
3. Kurtzke JF. Rating neurologic impairment in multiple sclerosis: an expanded disability status scale (EDSS). Neurology 1983 Nov;33(11):1444-1452. [Medline: 6685237]
4. Campbell JD, Ghushchyan V, Brett McQueen R, Cahoon-Metzger S, Livingston T, Vollmer T, et al. Burden of multiple sclerosis on direct, indirect costs and quality of life: National US estimates. Mult Scler Relat Disord 2014 Mar;3(2):227-236. [doi: 10.1016/j.msard.2013.09.004] [Medline: 25878010]

XSL•FO
RenderX

5. Reich DS, Lucchinetti CF, Calabresi PA. Multiple Sclerosis. N Engl J Med 2018 Jan 11;378(2):169-180 [FREE Full text] [doi: 10.1056/NEJMra1401483] [Medline: 29320652]

6. Munger KL, Zhang SM, O'Reilly E, Hernán MA, Olek MJ, Willett WC, et al. Vitamin D intake and incidence of multiple sclerosis. Neurology 2004 Jan 13;62(1):60-65. [doi: 10.1212/01.wnl.0000101723.79681.38] [Medline: 14718698]

7. Hämäläinen P, Ikonen A, Romberg A, Helenius H, Ruutiainen J. The effects of heat stress on cognition in persons with multiple sclerosis. Mult Scler 2012 Apr;18(4):489-497. [doi: 10.1177/1352458511422926] [Medline: 21914688]

8. Romberg A, Ikonen A, Ruutiainen J, Virtanen A, Hämäläinen P. The effects of heat stress on physical functioning in persons with multiple sclerosis. J Neurol Sci 2012 Aug 15;319(1-2):42-46. [doi: 10.1016/j.jns.2012.05.024] [Medline: 22647586]

9. Davis SL, Wilson TE, White AT, Frohman EM. Thermoregulation in multiple sclerosis. J Appl Physiol (1985) 2010 Nov;109(5):1531-1537. [doi: 10.1152/japplphysiol.00460.2010] [Medline: 20671034]

10. Mohr DC, Hart SL, Julian L, Cox D, Pelletier D. Association between stressful life events and exacerbation in multiple sclerosis: a meta-analysis. BMJ 2004 Mar 27;328(7442):731 [FREE Full text] [doi: 10.1136/bmj.38041.724421.55] [Medline: 15033880]

11. Marrie RA, Cohen J, Stuve O, Trojano M, Sørensen PS, Reingold S, et al. A systematic review of the incidence and prevalence of comorbidity in multiple sclerosis: overview. Mult Scler 2015 Mar;21(3):263-281 [FREE Full text] [doi: 10.1177/1352458514564491] [Medline: 25623244]

12. Capkun G, Dahlke F, Lahoz R, Nordstrom B, Tilson HH, Cutter G, et al. Mortality and comorbidities in patients with multiple sclerosis compared with a population without multiple sclerosis: An observational study using the US Department of Defense administrative claims database. Mult Scler Relat Disord 2015 Nov;4(6):546-554 [FREE Full text] [doi: 10.1016/j.msard.2015.08.005] [Medline: 26590661]

13. Lerdal A, Celius EG, Krupp L, Dahl AA. A prospective study of patterns of fatigue in multiple sclerosis. Eur J Neurol 2007 Dec;14(12):1338-1343. [doi: 10.1111/j.1468-1331.2007.01974.x] [Medline: 17903208]

14. Hassan E. Recall Bias can be a Threat to Retrospective and Prospective Research Designs. IJE 2006 Jan;3(2):1-7. [doi: 10.5580/2732]

15. McPhail S, Haines T. Response shift, recall bias and their effect on measuring change in health-related quality of life amongst older hospital patients. Health Qual Life Outcomes 2010 Jul 10;8:65 [FREE Full text] [doi: 10.1186/1477-7525-8-65] [Medline: 20618978]

16. Schmier JK, Halpern MT. Patient recall and recall bias of health state and health status. Expert Rev Pharmacoecon Outcomes Res 2004 Apr;4(2):159-163. [doi: 10.1586/14737167.4.2.159] [Medline: 19807519]

17. D'Amico E, Haase R, Ziemssen T. Review: Patient-reported outcomes in multiple sclerosis care. Mult Scler Relat Disord 2019 Aug;33:61-66. [doi: 10.1016/j.msard.2019.05.019] [Medline: 31154262]

18. Solomon AJ, Weinshenker BG. Misdiagnosis of multiple sclerosis: frequency, causes, effects, and prevention. Curr Neurol Neurosci Rep 2013 Dec;13(12):403. [doi: 10.1007/s11910-013-0403-y] [Medline: 24142849]

19. Birnbaum F, Lewis D, Rosen RK, Ranney ML. Patient engagement and the design of digital health. Acad Emerg Med 2015 Jun;22(6):754-756 [FREE Full text] [doi: 10.1111/acem.12692] [Medline: 25997375]

20. Jongen PJ, Ter Veen G, Lemmens W, Donders R, van Noort E, Zeinstra E. The Interactive Web-Based Program MSmonitor for Self-Management and Multidisciplinary Care in Persons With Multiple Sclerosis: Quasi-Experimental Study of Short-Term Effects on Patient Empowerment. J Med Internet Res 2020 Mar 09;22(3):e14297 [FREE Full text] [doi: 10.2196/14297] [Medline: 32149713]

21. Ziemssen T, Piani-Meier D, Bennett B, Johnson C, Tinsley K, Trigg A, et al. A Physician-Completed Digital Tool for Evaluating Disease Progression (Multiple Sclerosis Progression Discussion Tool): Validation Study. J Med Internet Res 2020 Feb 12;22(2):e16932 [FREE Full text] [doi: 10.2196/16932] [Medline: 32049062]

22. Giunti G, Kool J, Rivera RO, Dorronzoro ZE. Exploring the Specific Needs of Persons with Multiple Sclerosis for mHealth Solutions for Physical Activity: Mixed-Methods Study. JMIR Mhealth Uhealth 2018 Feb 09;6(2):e37 [FREE Full text] [doi: 10.2196/mhealth.8996] [Medline: 29426814]

23. Creagh AP, Simillion C, Scotland A, Lipsmeier F, Bernasconi C, Belachew S, et al. Smartphone-based remote assessment of upper extremity function for multiple sclerosis using the Draw a Shape Test. Physiol Meas 2020 Jun 19;41(5):054002. [doi: 10.1088/1361-6579/ab8771] [Medline: 32259798]

24. Ysrraelit MC, Fiol MP, Gaitán MI, Correale J. Quality of Life Assessment in Multiple Sclerosis: Different Perception between Patients and Neurologists. Front Neurol 2017;8:729 [FREE Full text] [doi: 10.3389/fneur.2017.00729] [Medline: 29375468]

25. Bot BM, Suver C, Neto EC, Kellen M, Klein A, Bare C, et al. The mPower study, Parkinson disease mobile data collected using ResearchKit. Sci Data 2016 Mar 03;3:160011 [FREE Full text] [doi: 10.1038/sdata.2016.11] [Medline: 26938265]

26. Chan YY, Wang P, Rogers L, Tignor N, Zweig M, Hershman SG, et al. The Asthma Mobile Health Study, a large-scale clinical observational study using ResearchKit. Nat Biotechnol 2017 Mar 13 [FREE Full text] [doi: 10.1038/nbt.3826] [Medline: 28288104]

27. Dorsey ER, Yvonne CY, McConnell MV, Shaw SY, Trister AD, Friend SH. The Use of Smartphones for Health Research. Acad Med 2017 Feb;92(2):157-160. [doi: 10.1097/ACM.0000000000001205] [Medline: 27119325]

28. Pratap A, Renn BN, Volponi J, Mooney SD, Gazzaley A, Arean PA, et al. Using Mobile Apps to Assess and Treat Depression in Hispanic and Latino Populations: Fully Remote Randomized Clinical Trial. J Med Internet Res 2018 Aug 09;20(8):e10130 [FREE Full text] [doi: 10.2196/10130] [Medline: 30093372]

29. Crouthamel M, Quattrocchi E, Watts S, Wang S, Berry P, Garcia-Gancedo L, et al. Using a ResearchKit Smartphone App to Collect Rheumatoid Arthritis Symptoms From Real-World Participants: Feasibility Study. JMIR Mhealth Uhealth 2018 Sep 13;6(9):e177. [doi: 10.2196/mhealth.9656] [Medline: 30213779]

30. McConnell MV, Shcherbina A, Pavlovic A, Homburger JR, Goldfeder RL, Waggot D, et al. Feasibility of Obtaining Measures of Lifestyle From a Smartphone App: The MyHeart Counts Cardiovascular Health Study. JAMA Cardiol 2017 Jan 01;2(1):67-76. [doi: 10.1001/jamacardio.2016.4395] [Medline: 27973671]

31. Bove R, White CC, Giovannoni G, Glanz B, Golubchikov V, Hujol J, et al. Evaluating more naturalistic outcome measures: A 1-year smartphone study in multiple sclerosis. Neurol Neuroimmunol Neuroinflamm 2015 Dec;2(6):e162 [FREE Full text] [doi: 10.1212/NXI.0000000000000162] [Medline: 26516627]

32. D'hooghe M, Van GG, Kos D, Bouquiaux O, Cambron M, Decoo D, et al. Improving fatigue in multiple sclerosis by smartphone-supported energy management: The MS TeleCoach feasibility study. Mult Scler Relat Disord 2018 May;22:90-96 [FREE Full text] [doi: 10.1016/j.msard.2018.03.020] [Medline: 29649789]

33. Genentech Inc. Floodlight Open. 2020. URL: https://floodlightopen.com/en-US/ [accessed 2020-10-08]

34. Midaglia L, Mulero P, Montalban X, Graves J, Hauser SL, Julian L, et al. Adherence and Satisfaction of Smartphone- and Smartwatch-Based Remote Active Testing and Passive Monitoring in People With Multiple Sclerosis: Nonrandomized Interventional Feasibility Study. J Med Internet Res 2019 Aug 30;21(8):e14863 [FREE Full text] [doi: 10.2196/14863] [Medline: 31471961]

35. Limmroth V, Hechenbichler K, Müller C, Schürks M. Assessment of Medication Adherence Using a Medical App Among Patients With Multiple Sclerosis Treated With Interferon Beta-1b: Pilot Digital Observational Study (PROmyBETAapp). J Med Internet Res 2019 Jul 29;21(7):e14373 [FREE Full text] [doi: 10.2196/14373] [Medline: 31359863]

36. elevateMS. Sage Bionetworks, Novartis Pharmaceuticals Corporation. URL: http://www.elevatems.org/ [accessed 2020-05-01]

37. Doerr M, Maguire Truong A, Bot BM, Wilbanks J, Suver C, Mangravite LM. Formative Evaluation of Participant Experience With Mobile eConsent in the App-Mediated Parkinson mPower Study: A Mixed Methods Study. JMIR Mhealth Uhealth 2017 Feb 16;5(2):e14 [FREE Full text] [doi: 10.2196/mhealth.6521] [Medline: 28209557]

38. Grayson S, Suver C, Wilbanks J, Doerr M. Open Data Sharing in the 21st Century: Sage Bionetworks' Qualified Research Program and Its Application in mHealth Data Release. Social Science Research Network. 2020. URL: https://ssrn.com/abstract=3502410 [accessed 2020-10-08]

39. Hohol MJ, Orav EJ, Weiner HL. Disease steps in multiple sclerosis: a simple approach to evaluate disease progression. Neurology 1995 Feb;45(2):251-255. [Medline: 7854521]

40. Cella D, Lai J, Nowinski CJ, Victorson D, Peterman A, Miller D, et al. Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. Neurology 2012 Jun 05;78(23):1860-1867 [FREE Full text] [doi: 10.1212/WNL.0b013e318258f744] [Medline: 22573626]

41. Cella D, Nowinski C, Peterman A, Victorson D, Miller D, Lai J, et al. The neurology quality-of-life measurement initiative. Arch Phys Med Rehabil 2011 Oct;92(10 Suppl):S28-S36 [FREE Full text] [doi: 10.1016/j.apmr.2011.01.025] [Medline: 21958920]

42. User Manual for the Quality of Life in Neurological Disorders (Neuro-QoL) Measures, Version 2. Bethesda, MD: National Institute of Neurological Disorders and Stroke (NINDS); 2015. URL: http://www.healthmeasures.net/images/neuro_qol/Neuro-QOL_User_Manual_v2_24Mar2015.pdf [accessed 2020-10-09]

43. Cook KF, Victorson DE, Cella D, Schalet BD, Miller D. Creating meaningful cut-scores for Neuro-QOL measures of fatigue, physical functioning, and sleep disturbance using standard setting with patients and providers. Qual Life Res 2015 Mar;24(3):575-589. [doi: 10.1007/s11136-014-0790-9] [Medline: 25148759]

44. Neuro-QoL™ Score Cut Points. 2020. URL: http://www.healthmeasures.net/score-and-interpret/interpret-scores/neuro-qol/neuro-qol-score-cut-points [accessed 2020-10-09]

45. Miller DM, Bethoux F, Victorson D, Nowinski CJ, Buono S, Lai J, et al. Validating Neuro-QoL short forms and targeted scales with people who have multiple sclerosis. Mult Scler 2016 May;22(6):830-841 [FREE Full text] [doi: 10.1177/1352458515599450] [Medline: 26238464]

46. Validated neurocognitive testing: Digit Symbol Substitution. Iowa City, IA: BrainBaseline; 2020. URL: https://www.brainbaseline.com/product/cognitive-testing#symbolsub [accessed 2020-05-01]

47. Snyder P, Tummalacherla M, Perumal T, Omberg L. mhealthtools: A Modular R Package for Extracting Features from Mobile and Wearable Sensor Data. J Open Source Software 2020;5(47):2106 [FREE Full text] [doi: 10.21105/joss.02106]

48. ElevateMS Feature Definitions. Seattle, WA: Sage Bionetworks; 2020. URL: https://github.com/Sage-Bionetworks/elevateMS_data_release/blob/master/FeatureDefinitions.md [accessed 2020-10-09]

49. Bridge Platform. Seattle, WA: Sage Bionetworks; 2020. URL: https://sagebionetworks.org/bridge-platform/ [accessed 2020-10-09]

50. Bates D, Mächler M, Bolker B, Walker S. Fitting Linear Mixed-Effects Models Using. J Stat Soft 2015;67(1). [doi: 10.18637/jss.v067.i01]

XSL•FO
RenderX

51. Kuznetsova A, Brockhoff PB, Christensen RHB. lmerTest Package: Tests in Linear Mixed Effects Models. J Stat Soft 2017;82(13). [doi: 10.18637/jss.v082.i13]

52. The R Project for Statistical Computing. Vienna, Austria: The R Foundation; 2018. URL: https://www.R-project.org/ [accessed 2020-10-09]

53. Supplementary Data ElevateMS Manuscript. 2020. URL: http://doi.org/10.7303/syn22005477 [accessed 2020-10-10]

54. Synapse. URL: http://www.synapse.org/elevateMS [accessed 2020-10-10]

55. Synapse - Baseline Characteristics vs Active Test Performance (syn22006660.2). Seattle, WA: Sage Bionetworks; 2020. URL: http://doi.org/10.7303/syn22006660.2 [accessed 2020-05-01]

56. Synapse - Neuro-QoL vs Active Test Performance (syn22008560.3). Seattle, WA: Sage Bionetworks; 2020. URL: http://doi.org/10.7303/syn22008560.3 [accessed 2020-05-01]

57. Synapse - Weather vs Active Test Performance (syn22033185.1). Seattle, WA: Sage Bionetworks; 2020. URL: https://doi.org/10.7303/syn22033185.1 [accessed 2020-05-01]

58. Synapse - Weather vs Daily Check-ins (syn22033192.2). Seattle, WA: Sage Bionetworks; 2020. URL: https://doi.org/10.7303/syn22033192.2 [accessed 2020-05-01]

59. Synapse - Weather vs Daily Check-ins + Sensitivity Analysis (syn22033187.2). Seattle, WA: Sage Bionetworks; 2020. URL: https://doi.org/10.7303/syn22033187.2 [accessed 2020-05-01]

60. Dilokthornsakul P, Valuck RJ, Nair KV, Corboy JR, Allen RR, Campbell JD. Multiple sclerosis prevalence in the United States commercially insured population. Neurology 2016 Mar 15;86(11):1014-1021 [FREE Full text] [doi: 10.1212/WNL.0000000000002469] [Medline: 26888980]

61. Harbo HF, Gold R, Tintoré M. Sex and gender issues in multiple sclerosis. Ther Adv Neurol Disord 2013 Jul;6(4):237-248 [FREE Full text] [doi: 10.1177/1756285613488434] [Medline: 23858327]

62. Pratap A, Neto EC, Snyder P, Stepnowsky C, Elhadad N, Grant D, et al. Indicators of retention in remote digital health studies: a cross-study evaluation of 100,000 participants. NPJ Digit Med 2020;3:21 [FREE Full text] [doi: 10.1038/s41746-020-0224-8] [Medline: 32128451]

63. Hirpara DH, Gupta V, Brown L, Kidane B. Patient-reported outcomes in lung and esophageal cancer. J Thorac Dis 2019 Mar;11(Suppl 4):S509-S514 [FREE Full text] [doi: 10.21037/jtd.2019.01.02] [Medline: 31032069]

64. Kalluri M, Luppi F, Ferrara G. What Patients With Idiopathic Pulmonary Fibrosis and Caregivers Want: Filling the Gaps With Patient Reported Outcomes and Experience Measures. Am J Med 2020 Mar;133(3):281-289 [FREE Full text] [doi: 10.1016/j.amjmed.2019.08.032] [Medline: 31520625]

65. Nair D, Wilson FP. Patient-Reported Outcome Measures for Adults With Kidney Disease: Current Measures, Ongoing Initiatives, and Future Opportunities for Incorporation Into Patient-Centered Kidney Care. Am J Kidney Dis 2019 Dec;74(6):791-802. [doi: 10.1053/j.ajkd.2019.05.025] [Medline: 31492487]

66. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ 2015 Feb 10;350:g7818. [doi: 10.1136/bmj.g7818] [Medline: 25670183]

67. Calvert M, Kyte D, Mercieca-Bebber R, Slade A, Chan A, King MT, the SPIRIT-PRO Group, et al. Guidelines for Inclusion of Patient-Reported Outcomes in Clinical Trial Protocols: The SPIRIT-PRO Extension. JAMA 2018 Feb 06;319(5):483-494. [doi: 10.1001/jama.2017.21903] [Medline: 29411037]

68. Santana M, Feeny D. Framework to assess the effects of using patient-reported outcome measures in chronic care management. Qual Life Res 2014 Jun;23(5):1505-1513. [doi: 10.1007/s11136-013-0596-1] [Medline: 24318085]

69. Appendix 2 to the Guideline on the Evaluation of Anticancer Medicinal Products in Man. Amsterdam, The Netherlands: European Medicines Agency; 2016. URL: https://www.ema.europa.eu/en/documents/other/appendix-2-guideline-evaluation-anticancer-medicinal-products-man_en.pdf [accessed 2020-10-09]

70. Value and Use of Patient-Reported Outcomes (PROs) in Assessing Effects of Medical Devices. Silver Spring, MD: US Food and Drug Administration; 2019. URL: https://www.fda.gov/media/109626/download [accessed 2020-10-09]

71. PRO Report Appendix: Patient-Reported Outcome Measure (PRO) Case Studies. Silver Spring, MD: US Food and Drug Administration; 2019. URL: https://www.fda.gov/media/125193/download [accessed 2020-10-09]

72. Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. Silver Spring, MD: US Food and Drug Administration; 2009. URL: https://www.fda.gov/regulatory-information/search-fda-guidance-documents/patient-reported-outcome-measures-use-medical-product-development-support-labeling-claims [accessed 2020-10-09]

73. Haase R, Schultheiss T, Kempcke R, Thomas K, Ziemssen T. Use and acceptance of electronic communication by patients with multiple sclerosis: a multicenter questionnaire study. J Med Internet Res 2012;14(5):e135 [FREE Full text] [doi: 10.2196/jmir.2133] [Medline: 23069209]

74. Kern R, Haase R, Eisele JC, Thomas K, Ziemssen T. Designing an Electronic Patient Management System for Multiple Sclerosis: Building a Next Generation Multiple Sclerosis Documentation System. Interact J Med Res 2016 Jan 08;5(1):e2 [FREE Full text] [doi: 10.2196/ijmr.4549] [Medline: 26746977]

75. Kremenchutzky M, Walt L. Perceptions of health status in multiple sclerosis patients and their doctors. Can J Neurol Sci 2013 Mar;40(2):210-218. [doi: 10.1017/s0317167100013755] [Medline: 23419570]

76. Stuchiner T, Chen C, Baraban E, Cohan S. The Pacific Northwest MS Registry: Year 4 Update and Diagnosis Validation (Poster DX82). 2012 Presented at: Consortium of the Multiple Sclerosis Centers (CMSC) Annual Meeting; May 31–June 4, 2012; San Diego, CA, USA.

77. Verdier-Taillefer MH, Roullet E, Cesaro P, Alpérovitch A. Validation of self-reported neurological disability in multiple sclerosis. Int J Epidemiol 1994 Feb;23(1):148-154. [doi: 10.1093/ije/23.1.148] [Medline: 8194911]

78. Frahm N, Hecker M, Zettl UK. Multi-drug use among patients with multiple sclerosis: A cross-sectional study of associations to clinicodemographic factors. Sci Rep 2019 Mar 06;9(1):3743 [FREE Full text] [doi: 10.1038/s41598-019-40283-5] [Medline: 30842515]

79. Jelinek GA, Weiland TJ, Hadgkiss EJ, Marck CH, Pereira N, van DMDM. Medication use in a large international sample of people with multiple sclerosis: associations with quality of life, relapse rate and disability. Neurol Res 2015 Aug;37(8):662-673 [FREE Full text] [doi: 10.1179/1743132815Y.0000000036] [Medline: 25905471]

80. FDA-NIH Biomarker Working Group. BEST (Biomarkers, EndpointS, and other Tools) Resource. Silver Spring, MD: US Food and Drug Administration; 2016.

81. Hunter SF. Overview and diagnosis of multiple sclerosis. Am J Manag Care 2016 Jun;22(6 Suppl):s141-s150 [FREE Full text] [Medline: 27356023]

Abbreviations

DMT: disease-modifying therapy
DSST: Digit Symbol Substitution Test
LME: linear mixed effects
MS: multiple sclerosis
Neuro-QoL: Quality of Life in Neurological Disorders
PDDS: Patient-Determined Disease Steps
PRO: patient-reported outcome
QoL: quality of life

XSL•FO
RenderX