

Validated, Quantitative, Machine Learning-Generated Neurologic Assessment of Multiple Sclerosis Using a Mobile Application

Sharon Stoll, DO; Charisse Litchman, MD; Noah Rubin, BS; Larry Rubin, MPH; and Timothy Vartanian MD, PhD

ABSTRACT

BACKGROUND: The BeCare MS Link mobile app collects data as users complete different in-app assessments. It was specifically developed to evaluate the symptomatology and neurologic function of patients with multiple sclerosis (MS) and to become a digital equivalent of the Expanded Disability Status Scale (EDSS) and other standard clinical metrics of MS progression.

METHODS: Our research compared EDSS scores derived from the BeCare MS link app to EDSS scores derived from neurologist assessment for the same cohort of 35 patients diagnosed with MS. App-derived data were supplied to 4 different machine learning algorithms (MLAs) with an independent EDSS score prediction generated from each. These scores were compared with the clinically derived EDSS score to assess the similarity of the scores and to determine an accuracy estimate for each.

RESULTS: Of the 4 MLAs employed, the most accurate MLA produced 19 EDSS score predictions that exactly matched the clinically derived scores, 21 score predictions within 0.5 EDSS points, and 32 score predictions within 1 EDSS point. The remaining MLAs also provided a relatively high level of accuracy in predicting EDSS scores when compared with clinically derived EDSS, with over 80% of scores predicted within 1 point and a mean squared error with a range of 1.05 to 1.37.

CONCLUSIONS: The BeCare MS Link app can replicate the clinically derived EDSS assessment of a patient with MS. The app may also offer a more complete evaluation of disability in patients with MS.

Int J MS Care. 2024;26(2):69-74. doi:10.7224/1537-2073.2023-009

Multiple sclerosis (MS) is a chronic neurological disorder affecting more than 2.5 million individuals globally and approximately 1 million individuals in the United States alone.^{1,2} Individuals with MS can experience a broad symptomatology including pain, fatigue, depression, memory loss, and difficulty ambulating. There are effective medications available and others are being developed to help to prevent disease progression and associated disability.³

To determine when treatment regimens must be modified to optimize care, physicians rely on patient-reported symptoms and on physical changes detected through clinical examination. There are standard disability scales used by MS specialists to help quantify disability progression and neurologic dysfunction.^{4,5} The gold standard of these metrics is the Expanded Disability Status Scale (EDSS),⁴ but this scale and other clinical assessments are time-consuming and cumbersome.⁶ Reliance on the currently available disability status scores, which have limited sensitivity, combined with infrequent in-person visits, often delay treatment modifications. This is because decline in neurologic function and cognition is not typically detected in patients at the first sign of progression.^{7,8}

During the COVID-19 pandemic, most routine and nonessential health care appointments were switched to telehealth or placed on hold.⁹ The pandemic has therefore revealed the potential for telemedicine in the care of numerous chronic conditions, such as MS, in order to protect vulnerable individuals from harm, reduce the burden of clinical visits, and improve the early detection of degenerative or progressive disease.

The BeCare MS Link app—compatible with both Android and Apple operating systems—was developed to evaluate the symptomatology and neurologic function of patients and act as a digital equivalent to the clinically derived EDSS (ie, EDSS scores as determined by neurologists). By logging into the app, a user establishes a session in which they perform certain app-directed assessments to evaluate their MS symptoms, and the resultant data are collected and analyzed by machine learning

From the Department of Neurology, Yale School of Medicine, Yale University, New Haven, CT (SS, CL); Yale New Haven Hospital, New Haven, CT (SS, CL); BeCare Link, Rumson, NJ (NR, LR); and Weill Cornell Medicine, New York, NY (TV). *Correspondence:* Sharon Stoll, DO, Department of Neurology, Yale School of Medicine, Yale University, New Haven, CT 06510; *email:* sharon.stoll@yale.edu.

Note: Supplementary material for this article is available online at ijmsc.org.

© 2024 Consortium of Multiple Sclerosis Centers.

algorithms (MLAs) to produce an app-derived EDSS score. The measurements and data obtained from these assessments correspond to the neurologist-determined EDSS, which is currently the most widely used MS assessment tool.⁴ Although the EDSS is primarily based on ambulation,⁸ the app-based assessments evaluate the functionality of a range of nervous system components, including pyramidal, cerebellar, sensory, visual, mental, and motor processes. The app also collects information on diet, mood, memory loss, missed workdays, limiting symptoms, demographics, comorbidities, disease severity, and prior MS treatments. The BeCare MS Link app, as well as being a digital equivalent to a clinically derived EDSS, may provide a more complete evaluation of disability in patients with MS.

The primary objective of this clinical trial was to validate the EDSS scores derived from the BeCare MS Link app compared with the clinically derived EDSS scores of the same patients, as determined by neurologists at the Weill Cornell Institute. The secondary objective was to generate an EDSS score for individual patients that mirrors the EDSS score obtained via neurologist assessment. These objectives differ as the former is a comparison of app versus clinical measurements, while the second measurement is to test the accuracy of the MLAs used by the app.

By using the BeCare MS Link app, patients can gain greater control over the management of their disease. Patients can take the tests on the BeCare MS Link app at home and as often as they choose, meaning that neurologic deterioration can be detected earlier, and that the effectiveness of drug therapy can be followed more quantifiably and longitudinally. Ultimately, this minimizes the need for higher levels of care because physicians may be alerted to problems well before the need for an emergency department visit or inpatient admission.

METHODS

In total, 35 patients were enrolled in the study over a period of 1 year from the time of institutional review board approval. Patients had the opportunity to complete the study up to 3 months after enrollment, but in most cases enrollment and study completion took place on the same day. Research coordinators at the clinical site screened eligible patients and obtained informed consent. The trial was registered on ClinicalTrials.gov as NCT04281160.

Inclusion and Exclusion Criteria

The inclusion criteria consisted of patients aged 18 to 75 years with clinically definite MS based on the revised McDonald criteria.¹⁰ Patients needed to have mild to moderate disability in 1 or more of the modalities assessed by the BeCare MS Link app and a baseline clinically derived EDSS score of 0.0 to 6.5. Patients had to be able to speak English and to provide informed consent. Patients were excluded if they had a clinically derived EDSS score higher than 6.5, were unable to provide consent, had congenital or traumatic loss of index finger or thumb, had impaired mobility or function owing to rheumatologic or other illnesses, or had neurologic impairment due to an illness other than MS.

Informed Consent

The details of the study protocol were reviewed with each patient and all questions were addressed. If a patient agreed to enroll in the study, an informed consent form was signed. The informed consent was kept in a study binder and a digital copy was made and stored on a secure server in compliance with the Health Insurance Portability and Accountability Act.

Assessments

At their clinical visit, the patients were evaluated by MS neurologists at the Weill Cornell Institute and a clinically-derived EDSS score was obtained. Additionally, patients were trained to use the BeCare MS Link app on their mobile devices and obtained their first set of scores by completing a sufficient number of app assessments. There were 11 assessments in the app: Arm Elevation, Path, Transcription Test, Contrast Sensitivity, Timed 25-Foot Walk (T25-FW), 6-Minute Walk, Time Up and Go (TUG), Tap Task, Stroop Test, Code Test and Memory Test. These assessments included measurements of motor mobility, fine motor function, upper extremity coordination, auditory comprehension, time to walk predetermined distances, visual acuity, visual tracking of objects, cognitive function, memory, and vibration sense (FIGURE S1). Gold standard measurements for determining mobility function in MS, the T25-FW and TUG, were compared with similar evaluation modalities via the app in order to assess the agreement between clinician-derived and app-derived measurements.¹¹ Additional diagnostic procedures measured by the app were compared with the current clinical gold standard for disability in MS, the Kurtzke EDSS.⁴ The app questionnaires also collected data on the duration and severity of disease, success of various therapies, patient demographics, quality of life, mood, dietary habits, bowel and bladder function, and disease comorbidity (FIGURE S2). Patients completed clinical assessments and the 11 app-based assessments in 1 visit. The app-based assessments were repeated 3 to 7 times to eliminate learning variation.

The app-based assessments were designed to have a unique sequence of steps that the user completes to successfully perform the assessment. Each assessment corresponds to a standard functional test that a clinician could perform during a neurological assessment. The app uses the accelerometer, gyroscope, or magnetic sensor in the patient's phone to determine when they take a step, make a transition to a different position or orientation, or move their arms.

Users only completed some of the assessments during a particular session and completed additional assessments during later sessions. After an assessment was completed, the data corresponding to the steps of the assessment were sent to the BeCare Link cloud, provided an internet connection from the mobile device to the BeCare Link cloud was available. After the assessment data were stored in the BeCare Link cloud, the MLAs could produce an app-derived EDSS score.

Four different MLAs—Huber regression, linear regression, random forest, and naïve Bayes—were used to analyze the app-derived data and calculate app-derived EDSS scores. Each MLA predicted an independent EDSS score. The app-derived

EDSS scores were then compared to the clinically derived EDSS score to determine whether the app could replicate the clinically derived EDSS scores, as well as to verify the accuracy of the app-derived EDSS scores.

Statistical Analysis

The analysis was based upon extracting certain key features from the app-derived raw data that were correlated with clinically obtained EDSS scores. A key feature was one that was an essential determinant of the EDSS score from the point of view of the MLAs.

Key features were generated from the raw data that modeled the state transitions associated with completing each assessment. For example, the TUG assessment was modeled as a sequence of state transitions: (1) sitting to standing, (2) walking, (3) turning around, (4) walking back to the chair, and (5) standing to sitting. Therefore, each assessment produced hundreds of data points that were used to calculate MLA-derived scores.

Another example is the T25-FW assessment for which the app-derived raw data were obtained from statistical measures such as the median time taken between each step and the aggregate variation. Multiple completed instances of the same chosen assessment were aggregated by taking order statistics for each feature, such as the median or 25th percentile. These order statistics became the feature values. This raw data had high dimensionality, which made it difficult to analyze, so principal component analysis (PCA) was used to reduce its dimensionality.¹²

The MLAs were initially trained using a set of control participants (ie, people who do not have MS and no evidence of disability). Using only these control participants, a z-score was computed for each feature value. Data points that exceeded 2 standard deviations from the control population mean were removed as statistical outliers. When collecting data from sensory inputs, it is standard to filter out results that are attributable to either noisy measurements or the participant misunderstanding the directions of the assessment. Using this filtered set of input data, the mean μ_0 and standard deviation σ_0 were computed, which allowed the computation of z-scores for each participant and feature value, using the standard calculation $(x-\mu_0)/\sigma_0$, where x is the participant's feature value. These z-scores will be referred to as the feature statistics.

The z-scores for certain feature values had small standard deviations owing to the limited range of data. For each participant, feature threshold tests were calculated, each of which was a binary variable: 1 if the corresponding feature statistic was considered evidence of disability and 0 otherwise.

Each feature was given a threshold value that determined whether a feature statistic was evidence of disease. For instance, a participant with a high feature statistic for the time between steps was assigned a 1 for the corresponding feature threshold test, given that difficulty walking is a symptom of disability. In other cases, low feature statistics were indicative of disability; for example, the coding assessment had a feature statistic corresponding to the number of numeric codes mapped correctly to icons.

PRACTICE POINTS



Assessment scores generated by the BeCare MS Link app include evaluations of arm movement, cerebellar function, cranial nerve function, and sensory function, as well as standard Expanded Disability Status Scale (EDSS) assessments, to potentially generate a more comprehensive assessment of multiple sclerosis disease progression than just the EDSS alone, which is heavily weighted for ambulation.

The app could become a valid digital equivalent to clinical assessments by neurologists, potentially allowing for greater data collection across more time points for improved disease monitoring, as well as improved convenience for patients. ■

Nevertheless, the results obtained from all models were used because the dispersion of results can be used to generate a confidence statistic on the accuracy of the reported score.

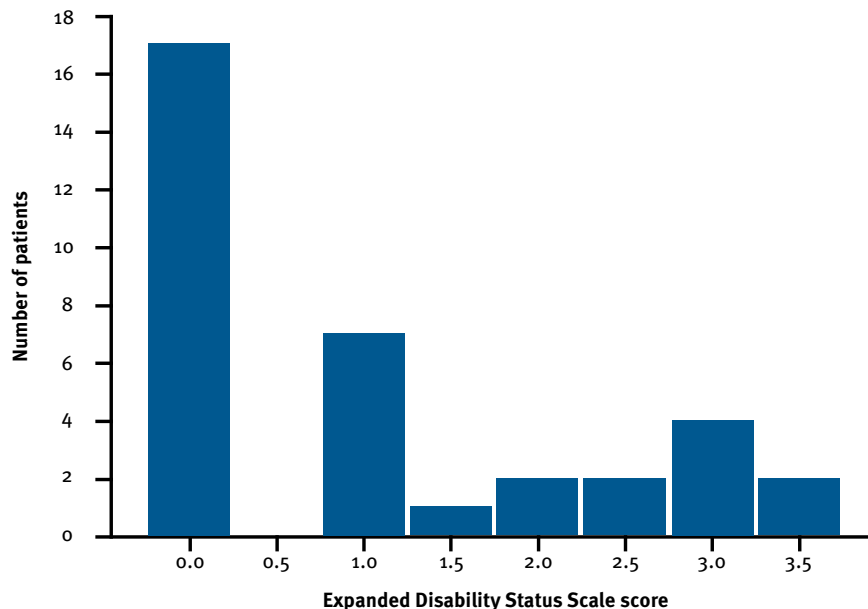
Given that other regression algorithms such as random forest and linear regression fail owing to missing data, feature threshold tests of 0 were substituted for those who had not completed the corresponding assessment. Essentially, participants were considered to have no disability unless there was evidence of it.

Feature Reduction

Overall, only 6 of the 11 BeCare MS Link app-based tasks were included in this analysis (see Results section). All together, these 6 assessments had 16 features, which were too many, relative to the number of eligible patients ($N = 35$), to perform direct regressions. The number of features is the same as the number of regression coefficients. When the number of regression coefficients is too large relative to the number of samples, the resulting model may become overfitted.

To avoid model overfitting, we sought to reduce the 16-feature threshold tests (regression coefficients) to a smaller number of features without losing too much information. We accomplished this in 2 different ways: We combined feature threshold tests according to (1) assessment and (2) functional system.

For (1), provided the corresponding assessment was completed, the mean value of each threshold test for each assessment was calculated, or was given a regressor value of 0 if the assessment was incomplete. This yielded 6 regressors.

FIGURE 1. Distribution of Clinically Derived Expanded Disability Status Scale Scores

For (2), assessments were grouped by the functional system they measured, allowing for overlap. For each functional system, the regressor was the mean of the means of each completed assessment associated with that functional system. A patient who did not perform any of the assessments associated with a functional system was given a regressor value of 0. Thus, this second method yielded 5 total regressors.

A machine-learning model to output the mean EDSS score was trained for each of these 2 sets of regressors. The model for each set of regressors used linear regression with the Huber loss function with a threshold parameter equal to 1.5 so that the regression was not dominated by outlier effects.¹³ The results of the models for both sets of regressors were then averaged to produce the final EDSS score for each independent MLA. The performance of the model was assessed with leave-one-out cross-validation, wherein a prediction for each patient was obtained by a model trained on all the other patients.¹⁴

RESULTS

Overall, 6 of the assessments had key features with high predictive value in determining EDSS scores: TUG, Tap Task, Path, T25-FW, Code Test, and the Contrast Sensitivity tests (TABLE S1). Although the clinical trial design was intended for all patients to perform the clinical assessments and the 11 BeCare MS Link app-based assessments, many patients were unable to stay for the time required to complete all app assessments; therefore, only 6 of the 11 BeCare MS Link app assessments were included in this analysis. The remaining 5 app assessments were excluded because they did not have sufficient participation by the cohort of patients.

To compare the app-derived EDSS score calculated from the 6 app-based assessments outlined in Table S1, the clinically

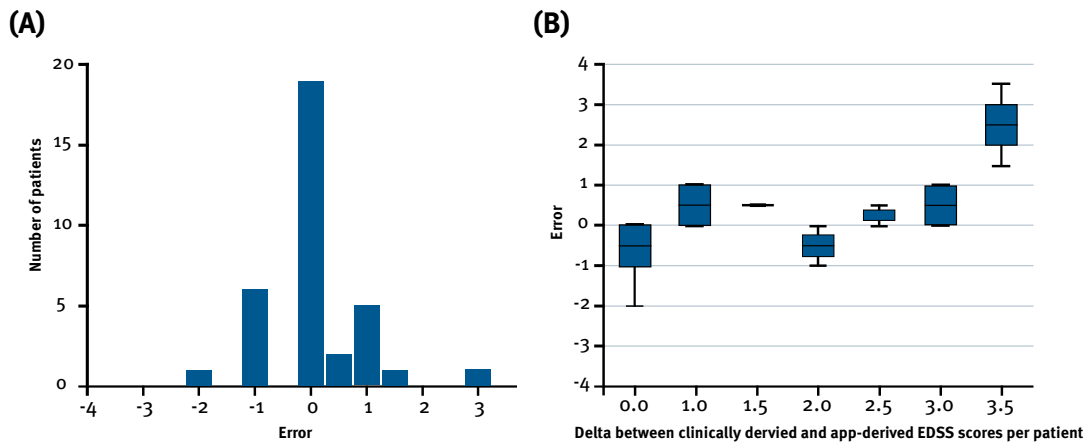
derived EDSS score was measured for all 35 participants. Their clinically derived EDSS scores ranged between 0 and 3.5 (FIGURE 1; TABLE S2), with 0.0 (48.6%) and 1.0 (20.0%) being the most common scores.

The clinically derived EDSS score for each patient was then compared to each of the app-derived EDSS scores that were independently calculated by the 4 different MLAs. The most accurate was the Huber regression. Nineteen patients (54.3%) had scores that were predicted exactly, 21 patients (60.0%) had scores that were predicted within half a point, and 32 patients (91.4%) had scores that were predicted within 1 point (FIGURES 2A and 2B; TABLE S3). The mean squared error of the Huber regression-derived EDSS score compared with the clinically derived EDSS score was 0.86 (Table S3).

The other 3 MLAs also provided a relatively high level of accuracy in predicting EDSS scores when compared with the clinically derived EDSS scores with more than 80% of scores predicted within 1 point and a mean squared error in the range of 1.05 to 1.37 (Table S3).

DISCUSSION

There are many benefits to using mobile apps to monitor patient health outside the clinic setting, and the COVID-19 pandemic has highlighted this.¹⁵ In the literature, studies have analyzed the use of mobile apps to monitor chronic conditions including congestive heart failure, diabetes, and MS.¹⁶⁻¹⁹ To the best of our knowledge, there is yet to be a mobile app capable of evaluating the neurological and physical functioning of patients. Other apps designed to monitor MS rely on patient-reported symptoms and they function more specifically as treatment trackers.^{20,21} The BeCare MS Link app is novel in its ability to measure neurologic dysfunction objectively, rather than relying on

FIGURE 2. Distribution of Prediction Errors

EDSS, Expanded Disability Status Scale; MLA, machine learning algorithm.

(A) Distribution of error for the app-derived EDSS score as predicted by the Huber regression MLA. The error of score 0 means an exact agreement of Huber regression-based app-derived EDSS and clinical EDSS scores. The other bars demonstrate the number of EDSS point differences, with +1 meaning an app-derived EDSS of 1 point higher than the clinically derived EDSS and -1 meaning an app-derived EDSS of 1 point lower than the clinically derived EDSS.

(B) Error distribution based on the clinically derived EDSS score. This graph demonstrates the difference between the app-derived EDSS and the clinically derived EDSS per patient.

patient-reported symptoms, and in its use of MLAs to calculate EDSS scores that can mimic clinically derived EDSS scores. Additionally, the BeCare MS Link can reduce the problematic inter-rater variation between 2 clinician assessments of the same patient,⁸ benefiting outcomes in clinical care as well as in clinical trials.

Patient access to quantifiable assessments that can be performed at home via a mobile app will result in a much greater frequency of data collection. The BeCare MS Link's MLAs can detect subtle changes in patient performance so that their treating physicians can intervene earlier than with the data derived from standard clinical scoring, which is collected on an infrequent basis. This is a marked improvement in patient monitoring of MS, which has a relapsing-remitting course. As the BeCare MS Link measurements are made every time the app is used, it is possible to study the relationship between real-time assessments, such as medication changes, dietary changes, and rehabilitation, with changes in the BeCare MS Link-derived EDSS score over time.

The data collected by the BeCare MS Link app are not only available to the patient and their physician but can also be used by the MLAs as training in order to improve the app's output in future EDSS calculations. The MLAs learn by updating their internal state each time they are presented with EDSS-related data. This way the accuracy and reproducibility of the app-derived EDSS scores are intended to improve with time, reducing the variability in EDSS scores that is often seen in the clinic. This analysis used MLAs because of the availability of neurologist-supplied EDSS scores that could be used in the future to further train the algorithms, thereby potentially minimizing computed output errors.^{15,22} Historically, assessment of disease progression has been based on the degree of impairment as compared with either a

population standard or the patient's historical performance. This technique suffers from a learning effect as patient scores improve after they become accustomed to the testing.²³ We have therefore trained our MLAs to account for this "drift" in the data owing to familiarity.

The app-derived EDSS scores can both objectively validate the patient report of clinical improvement or disability progression and help to discern true improvement from both placebo and learning effect. Changes in app-derived EDSS scores can then be reported to the treating physician who can determine whether an earlier follow-up might be beneficial. It would also be possible for investigators, insurers, and health care systems to study the large and comprehensive collection of data to assess comorbidities and the success of new and established treatments and rehabilitation techniques. Ultimately, because of the wide breadth of data collected by the app, it will be possible to learn much more about MS.

This study does have limitations. Despite confirming the equivalence of clinically derived and app-derived EDSS scores, a number of app-based assessments were excluded from the final analysis because many patients did not complete app-based assessments due to time constraints. In future studies, we hope to perform further analyses on the utility of the BeCare MS Link app, which will include these assessments, and we hope to categorically determine whether the app-based assessments provide a more rounded and thorough evaluation of MS than the clinician-derived EDSS score. Other limitations include the small size of the cohort and that the trial was conducted at a single site. Finally, the patients in this trial did not have EDSS scores greater than 3.5, so future work should focus on and include patients with higher levels of clinician-assessed disability. We hope to

expand the size of the cohort to include patients with a range of disability levels and conduct a multinational and multi-centered trial to fully characterize the accuracy and utility of the BeCare MS Link app as a digital equivalent, or potential replacement, of the clinician-derived EDSS score.

CONCLUSIONS

The BeCare MS Link app can replicate the assessment of a patient with MS by a clinically derived EDSS. The goal of the BeCare MS Link app is to improve the well-being of people with MS through the accurate assessment of neurologic function, which can be used to direct treatment decisions by the patients themselves as well as by their health care providers. The clinically derived EDSS score is heavily weighted for ambulation whereas assessments made by the app via MLAs include evaluations of arm movement, cerebellar function, cranial nerve function, and sensory function, offering a more complete evaluation of disability in patients with MS. These results show that the BeCare MS Link app has utility for both physician and patient to monitor stability, improvement, or progression and support the use of the app in clinical trials to acquire significantly more information at a greater number of time points with marked reduction in cost and reduced inter-rater variability in multicenter trials. ■

ACKNOWLEDGMENTS: Medical writing support was provided by Mark Elms, PhD, of PharmaGenesis London, London, UK, and Adeline Rosenberg, MSc, of Oxford PharmaGenesis, Oxford, UK, and was funded by BeCare Link, LLC.

FINANCIAL DISCLOSURES: Dr Stoll has served on scientific advisory boards for Bristol Myers Squibb; F. Hoffmann-La Roche Ltd; Forepoint Capital Partners; Genentech, Inc; Horizon Therapeutics, Inc; and TG Therapeutics, Inc; received research support from BeCare Link, LLC and MedDay Pharmaceuticals SA; has received compensation for consulting services, served on scientific advisory boards, and received speaker honorarium for Alexion Pharmaceuticals, Inc; Biogen, Inc; Bristol Myers Squibb; EMD Serono, Inc; Horizon Therapeutics, Inc; Novartis AG; Roche's Genentech, Inc; and Sanofi Genzyme; is CEO of Global Consultant MD; and has served on the steering committee of Horizon Therapeutics, Inc and Roche's Genentech, Inc. Dr Lichtman has received personal compensation for consulting, serving on a scientific advisory board, speaking, or other activities with Amgen, Inc; holds stock and/or stock options in BeCare Link, LLC, which sponsored this research; and has received personal compensation for consulting, serving on a scientific advisory board, speaking, or other activities with Amgen, Inc. Dr Noah Rubin has no conflicts of interest to declare. Dr Larry Rubin has received personal compensation for consulting, serving on a scientific advisory board, speaking, or other activities with BeCare Link, LLC; has received compensation for serving on the board of directors of BeCare Link, LLC; and has received royalty, license fees, or contractual rights payments from BeCare Link, LLC. Dr Vartanian has received personal compensation for consulting, serving on a scientific advisory board, speaking, or

other activities with Biogen, Inc; Novartis AG; Roche's Genentech, Inc; and Sanofi Genzyme.

FUNDING/SUPPORT: Funding for this study was provided by BeCare Link, LLC.

REFERENCES

- Wallin MT, Culpepper WJ, Campbell JD, et al. The prevalence of MS in the United States: a population-based estimate using health claims data. *Neurology*. 2019;92(10):e1029-1040. doi:10.1212/WNL.0000000000007035
- Walton C, King R, Rechtman L, et al. Rising prevalence of multiple sclerosis worldwide: insights from the Atlas of MS, third edition. *Mult Scler*. 2020;26(14):1816-1821. doi:10.1177/1352458520970841
- Dobson R, Giovannoni G. Multiple sclerosis—a review. *Eur J Neurol*. 2019;26(1):27-40. doi:10.1111/ene.13819
- Kurtzke JF. Disability rating scales in multiple sclerosis. *Ann N Y Acad Sci*. 1984;436:347-360. doi:10.1111/j.1749-6632.1984.tb14805.x
- Krupp LB, LaRocca NG, Muir-Nash J, Steinberg AD. The fatigue severity scale. Application to patients with multiple sclerosis and systemic lupus erythematosus. *Arch Neurol*. 1989;46(10):1121-1123. doi:10.1001/archneur.1989.00520460115022
- Alves P, Green E, Leavy M, et al. Validation of a machine learning approach to estimate expanded disability status scale scores for multiple sclerosis. *Mult Scler J Exp Transl Clin*. 2022;8(2):20552173221108635. doi:10.1177/20552173221108635
- Bethoux F, Bennett S. Evaluating walking in patients with multiple sclerosis: which assessment tools are useful in clinical practice? *Int J MS Care*. 2011;13(1):4-14. doi:10.7224/1537-2073-13.1.4
- Van Munster CEP, Uitdehaag B. Outcome measures in clinical trials for multiple sclerosis. *CNS Drugs*. 2017;31(3):217-236. doi:10.1007/s40263-017-0412-5
- Monaghesh E, Hajizadeh A. The role of telehealth during COVID-19 outbreak: a systematic review based on current evidence. *BMC Public Health*. 2020;20(1):1193. doi:10.1186/s12889-020-09301-4
- Polman CH, Reingold SC, Banwell B, et al. Diagnostic criteria for multiple sclerosis: 2010 revisions to the McDonald criteria. *Ann Neurol*. 2011;69(2):292-302. doi:10.1002/ana.22366
- Fritz NE, Eloyan A, Baynes M, Newsome SD, Calabresi PA, Zackowski KM. Distinguishing among multiple sclerosis fallers, near-fallers and non-fallers. *Mult Scler Relat Disord*. 2018;19:99-104. doi:10.1016/j.msard.2017.11.019
- Jolliffe IT, Cadima J. Principal component analysis: a review and recent developments. *Philos Trans A Math Phys Eng Sci*. 2016;374(2065):20150202. doi:10.1098/rsta.2015.0202
- Huang S, Wu Q. Robust pairwise learning with Huber loss. *J Complex*. 2021;66:101570. doi:10.1016/j.jco.2021.101570
- Cheng H, Garrick DJ, Fernando RL. Efficient strategies for leave-one-out cross validation for genomic best linear unbiased prediction. *J Anim Sci Biotechnol*. 2017;8:38. doi:10.1186/s40104-017-0164-6
- Ahmed Z, Mohamed K, Zeeshan S, Dong XQ. Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database (Oxford)*. 2020;2020:baaa010. doi:10.1093/database/baaa010
- Rudick RA, Miller D, Bethoux F, et al. The Multiple Sclerosis Performance Test (MSPT): an iPad-based disability assessment tool. *J Vis Exp*. 2014;88:e51318. doi:10.37971/51318
- Macaron G, Moss BP, Li H, et al. Technology-enabled assessments to enhance multiple sclerosis clinical care and research. *Neurol Clin Pract*. 2020;10(3):222-231. doi:10.1212/CPJ.0000000000000710
- Lee EY, Cha SA, Yun JS, et al. Efficacy of personalized diabetes self-care using an electronic medical record-integrated mobile app in patients with type 2 diabetes: 6-month randomized controlled trial. *J Med Internet Res*. 2022;24(7):e37430. doi:10.2196/37430
- Rao VN, Kaltenbach LA, Granger BB, et al. The association of digital health application use with heart failure care and outcomes: insights from CONNECT-HF. *J Card Fail*. 2022;28(10):1487-1496. doi:10.1016/j.cardfail.2022.07.050
- Balakrishnan P, Groenberg J, Jacshyn-Owen E, et al. Demographic patterns of MS patients using BRISA: an MS-specific app in Germany. *J Pers Med*. 2022;12(7):1100. doi:10.3390/jpm12071100
- Wendrich K, Krabbenborg L. Digital self-monitoring of multiple sclerosis: interview study with Dutch health care providers on the expected new configuration of roles and responsibilities. *JMIR mHealth uHealth*. 2022;10(4):e30224. doi:10.2196/30224
- Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol*. 2019;19(1):64. doi:10.1186/s12874-019-0681-4
- Woelfle T, Pless S, Wiencierz A, Kappos L, Naegelin Y, Lorscheider J. Practice effects of mobile tests of cognition, dexterity, and mobility on patients with multiple sclerosis: data analysis of a smartphone-based observational study. *J Med Internet Res*. 2021;23(11):e30394. doi:10.2196/30394