Characterizing Protein and MSDA Score Correlations and Minimum Detectable Change Amongst Stable Participants from the Measuring Outcomes and Value: an Integrated, Novel Solution for Generating insights in MS (MOVING MS) Study

Shannon McCurdy[1], Elisa Sheng[1], Kian Jalaleddini[1,2], Gargi Datta[1], Tamara Shabi[3], Annalise E. Miner[4], Alexander Hari[1], Srushti Tiwari[1], Wayne Hu[1], Franklin Faust[1], Kelly Leyden[1], Sarah Eagleman[1], David Brazel[1], Terrie Livingston[1], Ferhan Qureshi[1], Jennifer S. Graves[3]. [1] Octave Bioscience Inc., Menlo Park, USA. [2] Tigro, San Diego, CA, USA. [3] Department of Neurosciences, University of California San Diego, La Jolla, CA, USA. [4] Boston University CTE Center, Boston University Chobanian & Avedisian School of Medicine, Boston, MA, USA.

P1544

INTRODUCTION

The MOVING MS study was designed to evaluate the impact of the Octave digital care platform on multiple sclerosis (MS) care, with primary endpoints focused on unplanned healthcare utilization and patient and physician satisfaction with the digital care platform [1]. As an exploratory endpoint, serial blood draws were collected to measure the Multiple Sclerosis Disease Activity (MSDA) test– a validated 18 protein biomarker panel that generates an age- and sex-corrected Disease Activity (DA) score that has been validated using MRI and clinical assessments of disease activity [2, 3].

OBJECTIVES

- Evaluate the ability of the DA score to capture clinical stability in individuals with relapsing-remitting multiple sclerosis (RRMS) in a post-hoc analysis.
- Compare the DA score's performance to neurofilament light chain (NfL) and glial fibrillary acidic protein (GFAP) concentrations.
- Utilize the DA score's test-retest reliability estimates to characterize Minimal Detectable Change (MDC_{95}), which reflects the smallest individual-level change that exceeds expected measurement error with 95% confidence.

METHODS

Study Population: Clinically stable relapsing-remitting multiple sclerosis (RRMS) participants from the 12-month MOVING MS pilot study longitudinal cohort.

Inclusion Criteria:

- Completion of ≥2 MSDA Tests.
- Clinical Stability Criteria over the study period (see Figure 1).

MS Disease Activity (MSDA) Test [2, 3]:

- DA scores (1-10, 0.5 increments) derived from absolute concentrations (pg/mL) of 18 proteins (including NfL and GFAP), age, and sex.
- NfL log-transformed protein concentration (pg/mL).
- GFAP log-transformed protein concentration (pg/mL).

Statistical Analysis:

- **Test-retest Reliability:** Pearson correlation coefficient (r) between the first two successively collected DA scores. Cluster bootstrapped 95% Confidence Intervals (CIs) [4]. Compared to the test-retest reliability of single log-transformed protein concentrations.
- **Minimum Detectable Change (MDC_{95}):** Calculated at 95% confidence with cluster bootstrapped 95% CIs. MDC_{95}= 1.96 × sqrt(2) × s × sqrt(1-r), where s is the standard deviation of baseline scores and r is the test-retest reliability [5].
- **Longitudinal Change:** Average annual DA score change estimated using a random slope and intercept linear mixed model, adjusted for age and sex.

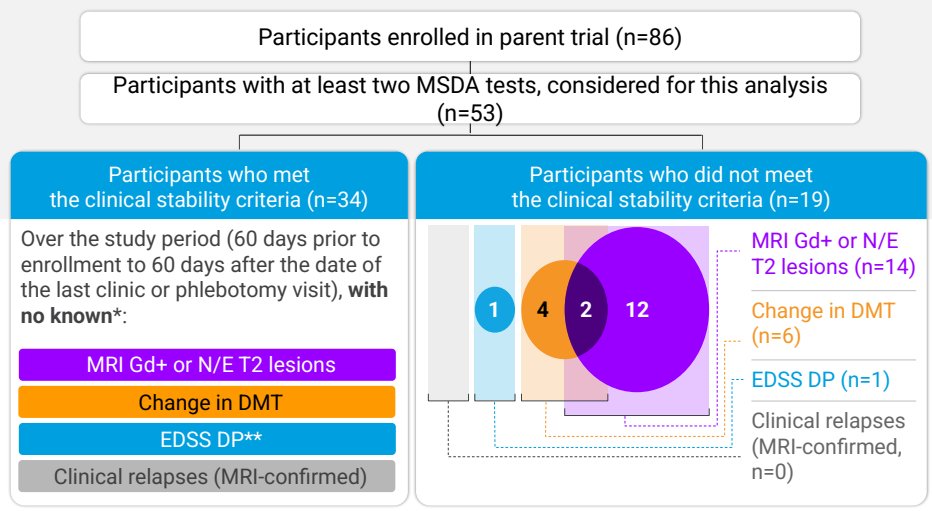
Statistic	Value	Overall
n		34
Age at baseline (years), mean (SD)		41.5 (10.5)
Sex, n (%)	F	27 (79.4)
Sex, n (%)	M	7 (20.6)
Disease duration at baseline (years), mean (SD)		8.1 (7.4)
Most recent EDSS at baseline*, mean (SD)		2.5 (1.4)
Disease modifying therapy at baseline, n (%)	Dimethyl Fumarate	1 (2.9)
	Diroximel Fumarate	1 (2.9)
	Fingolimod	3 (8.8)
	Glatiramer Acetate	6 (17.6)
	Natalizumab	2 (5.9)
	Ocrelizumab	11 (32.4)
	Ofatumumab	5 (14.7)
	Rituximab	3 (8.8)
	Teriflunomide	2 (5.9)
Intervention arm, n (%)	Case	17 (50.0)
	Wait List Control	17 (50.0)
Days between first and second MSDA timepoint, mean (SD)		132.9 (69.2)
DA score, mean (SD)		2.6 (1.4)
NfL log (pg/mL), mean (SD)		0.76 (0.16)
GFAP log (pg/mL), mean (SD)		1.98 (0.18)

Table 1: *Demographic characteristics of N=34 selected participants. *4 missing observations.*

KEY RESULTS

Of 86 original participants, 34 (108 samples) were determined to exhibit clinical stability over the study period (see Figure 1). See Table 1 for participant characteristics.

- The DA score had mean (standard deviation) of 2.71 (1.56) score units at the first timepoint (see Table 1).
- DA scores remained relatively stable and were predominantly low (see Figure 2).
- The DA score demonstrated strong test-retest reliability between the first two measurements (r=0.92, 95%CI=[0.79, 0.97]), while NfL (r=0.75, 95%CI=[0.47,0.91]) and GFAP (r=0.71, 95%CI=[0.46, 0.89]) showed lower reliability.
- The MDC_{95} for the DA score was 1.24 score units (95%CI=[0.74, 1.73]).
- On average, DA scores minimally changed over time, with a mean decrease of 0.47 score units/year (95%CI=[0.16, 0.79]).



*Absence of these events in medical records was interpreted as clinical stability.
**DP over the study period was: (1) ≥2-point EDSS increase if baseline EDSS = 0; (2) ≥1.5-point increase if baseline EDSS = 0.5; (3) ≥1-point increase if baseline EDSS between 1.0 and 5.0; and (4) ≥0.5-point increase if baseline EDSS ≥5.5.
DMT = Disease Modifying Therapy; DP = Disability Progression; Gd+ = gadolinium-enhancing; N/E=New/Enlarging.

Figure 1: *Flow diagram of participant selection.*

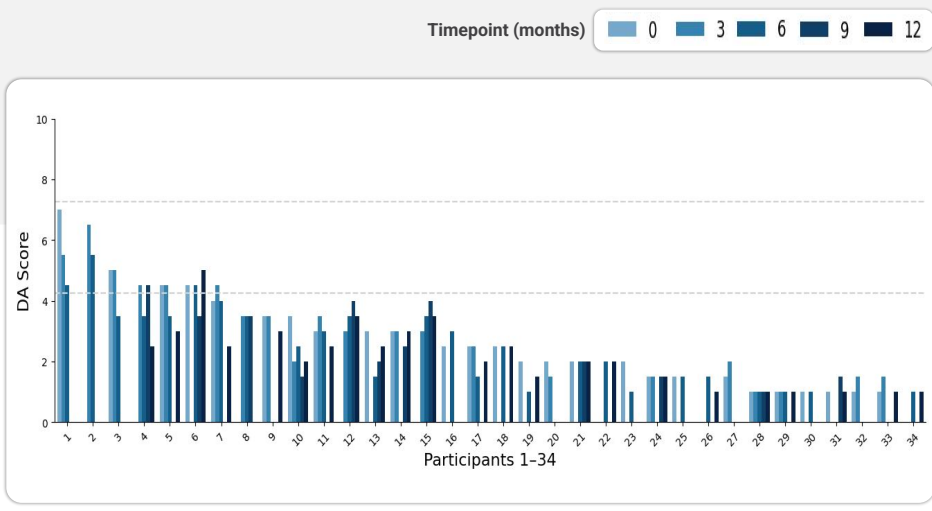


Figure 2: *Disease Activity (DA) scores for each participant over the study period. Horizontal lines indicate the DA score thresholds distinguishing Low (1.0 - 4.0) from Moderate (4.5 - 7.0) disease activity and Moderate from High (7.5 -10.0) disease activity.*

DISCUSSION

Strengths

Real-World Evidence: Data was sourced from a prospective clinical trial in a real-world clinical setting, enhancing the relevance of the findings.

Rigorous Analysis: Classical true-score theory was used to evaluate the DA score's reliability and define a threshold for meaningful change (MDC_{95}).

Limitations

Conservative Approach: Missing data was treated as stable, a necessary choice for this real-world study that may overestimate the MDC_{95}.

Pilot Study: Results are promising but require confirmation in larger, more definitive trials.

CONCLUSIONS

For most clinically stable participants, the MSDA score remained low and stable. The MSDA score directionally outperformed NfL and GFAP in test-retest reliability. The MDC_{95} provides a practical threshold for interpreting DA score changes: individual-level changes over 1.24 can be interpreted as reflecting a true change in disease activity, while smaller changes fall within expected measurement error and should be interpreted with caution. These findings support the DA score as a reliable biomarker for monitoring disease activity in RRMS.

Disclosure: S. McCurdy, E. Sheng, G. Datta, A. Hari, S. Tiwari, W. Hu, F. Faust, K. Leyden, S. Eagleman, D. Brazel, T. Livingston, and **F. Qureshi** are employees of Octave Bioscience. **K. Jalaleddini** reports consulting fees from Octave Bioscience, Inc. via Tigro LLC. He is currently employed at Komodo Health, which was not involved in the work. **T. Shabi** and **A. Miner** have no potential conflicts of interest. **J. Graves** has received research grants from Sanofi, Genentech, Ad Scientium, Octave and EMD Serono. She has received consulting fees from Octave, TR1X and Google.

References: [1] Miner A, Becich M, Gehman V, et al. Improving patient outcomes through a comprehensive care management platform (MOVING MS). ACTRIMS-ECTRIMS Meeting MS Virtual, 2020. [2] Qureshi F, Hu W, Loh L, et al. Analytical validation of a multi-protein, serum-based assay for disease activity assessments in multiple sclerosis. Proteomics Clin Appl 2023; 17: e2200018. [3] Chitnis T, Foley J, Ionete C, et al. Clinical validation of a multi-protein, serum-based assay for disease activity assessments in multiple sclerosis. Clin Immunol 2023; 253: 109688. [4] Huang FL. Using cluster bootstrapping to analyze nested data with a few clusters. Educ Psychol Meas 2018; 78: 297–318. [5] Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice, Third Edition. F.A. Davis Company, 2015.