Analytical validation of a multivariate proteomic serum-based assay for disease activity assessments in multiple sclerosis

Ferhan Qureshi,[1]* Wayne Hu,[1]* Louisa Loh,[1] Hemali Patel,[1] Maaria DeGuzman,[1] Michael Becich,[1†] Fatima Rubio da Costa,[1] Victor Gehman,[1] Fujun Zhang,[1] John Foley,[2] Tanuja Chitnis[3]

[1]Octave Bioscience, Inc., Menlo Park, CA, USA

[2]Rocky Mountain Multiple Sclerosis Clinic, Salt Lake City, UT, USA

[3]Brigham and Women's Hospital, Harvard Medical School, Boston, MA, USA

*These authors contributed equally to this work

[†]Employee of Octave Bioscience, Inc. at the time the study was completed

Correspondence

Ferhan Qureshi

1440 Obrien Drive, Suite B

Menlo Park, CA 94025, USA

Email: fqureshi@octavebio.com

Phone: 650-224-3072

Fax: 650-479-8876

Abbreviations:

%CV, percent coefficient of variation; APLPL1, amyloid beta precursor-like protein 1; CCL20, C-C motif chemokine ligand 20; CD6, cluster of differentiation 6; CDCP1, CUB domain-containing protein 1; C_{max}, maximum concentration; CNS, central nervous system, CNTN2, contactin 2; COL4A1, collagen type IV alpha-1; conc, concentration; CSF, cerebrospinal fluid; CXCL9, chemokine (C-X-C motif) ligand 9 (MIG); CXCL13, chemokine (C-X-C motif) ligand 13; DMT, disease-modifying therapy; DNA, deoxyribonucleic acid; FLRT2, fibronectin leucine-rich repeat transmembrane protein; GFAP, glial fibrillary acidic protein; GH, growth hormone; HAMA, human anti-mouse antibodies; HCl, hydrochloride; IL-12β, interleukin-12 subunit beta; LLOQ, lower limit of quantitation; LOQ, limit of quantitation; mAb, monoclonal antibody; Max, maximum; Min, minimum; MOG, myelin oligodendrocyte

glycoprotein; MRI, magnetic resonance imaging; MS, multiple sclerosis; MSDA, Multiple Sclerosis Disease Activity; Na, sodium; NfL, neurofilament light chain; OPG, osteoprotegerin; OPN, osteopontin; PCR, polymerase chain reaction; PEA, Proximity Extension Assay; PRTG, protogenin; prob, probability; qPCR, quantitative polymerase chain reaction; R^2, coefficient of determination; RA, rheumatoid arthritis; RF, rheumatoid factor; SD, standard deviation; SERPINA9, serpin family A member 9; TNFRSF10A, tumor necrosis factor receptor superfamily member 10A (TRAIL-R1); TNFSF13B, tumor necrosis factor superfamily member 13B (BAFF); ULOQ, upper limit of quantitation; VCAN, versican core protein

Key words:

Analytical characterization, analytical validation, biomarker, multiple sclerosis, proximity extension assay

Total word count: 4983

Running head: Analytical Validation of a Serum-Based Assay for Disease Activity Assessments in Multiple Sclerosis

Previous presentation: Part of this work, namely analytical validation of the individual biomarkers in the MSDA Test, was previously presented at the ACTRIMS 2021 Forum, Virtual (February 25–27, 2021)

Abstract

Purpose: To characterize and analytically validate the MSDA Test, a serum-based multiplex protein biomarker assay developed using Olink® PEA methodology.

Experimental design: Two lots of the MSDA Test panel were manufactured and subjected to a comprehensive analytical characterization and validation protocol to detect biomarkers present in the serum of patients with MS. Biomarker concentrations were incorporated into a final algorithm used for calculating four Disease Pathway scores (Immunomodulation, Neuroinflammation, Myelin Biology and Neuroaxonal Integrity) and an overall Disease Activity score.

Results: Analytical characterization demonstrated that the multiplex panel satisfied the criteria necessary for a fit-for-purpose validation considering the assay's intended clinical use. This panel met acceptability criteria for 18 biomarkers included in the final algorithm out of 21 biomarkers evaluated. VCAN was omitted based on factors outside of analytical validation; COL4A1 and GH were excluded based on imprecision and diurnal variability, respectively. Performance of the four Disease Pathway and overall Disease Activity scores met the established acceptability criteria.

Conclusions and clinical relevance: Analytical validation of this serum-based proteomic multiplex assay is the first step in establishing its potential utility as a quantitative, minimally invasive, and scalable biomarker panel to enhance the standard of care for patients with MS.

1 INTRODUCTION

Multiple Sclerosis (MS) is a chronic, neurodegenerative, immune-mediated disease of the CNS, characterized by inflammatory demyelination and neuronal damage.[1,2] MS has a complex disease course with variable symptoms or manifestations that can range from mild and self-limiting to severe.[1] The clinical course, after the first clinical manifestation of the disease, or clinically isolated syndrome, can vary.[3] The damage caused by MS typically leads to relapses, or acute attack of symptoms, followed by progressive disease.[4] Most treatments are effective in early stages of the relapsing/remitting form of the disease;[4,5] however, a delay in treatment can lead to irreversible damage.[6] Studies show that the extent of remyelination in early MS is greater than in chronic MS.[7] Clinical studies are underway to explore treatments targeting remyelination, which may slow or offset disease progression[8].

The McDonald Criteria, designed to improve the accuracy of MS diagnosis, established the use of MRI to show the accrual of lesions over time and space.[9] The revised McDonald Criteria substituted CSF oligoclonal immunoglobulin G bands for the second clinical/MRI finding.[10] Nonetheless, use of any of these assessments do not always accurately predict disease activity, course, progression, recurrence, or response to treatment.[11-13] As such, there is an unmet clinical need for objective and quantitative measures that can accurately diagnose MS, monitor disease activity, and promote individualized disease management.[13,14]

One major area of focus in MS is the identification of biomarkers in biological fluids, such as CSF or blood, to track pathogenesis, disease activity, and progression.[14,15] One of the key therapeutic strategies in MS is to reduce relapse, lesions, and brain atrophy at all disease stages.[4] As a result, new biomarkers for early MS diagnosis and disease activity monitoring can lead to prevention of disease progression, potentially reducing the patient's level of disease worsening.[14] The

dynamic range of proteins in CSF presents challenges when differentiating small disease-specific changes from inherent interindividual differences, especially as it relates to methodological variations.[16,17] CSF collection also requires invasive procedures, such as lumbar puncture. On the other hand, blood-based collection of biomarkers allows for safe, quick, and easy collection.[14] With these considerations, detection of biomarkers in blood is a viable and attractive option for the accurate diagnosis and assessment of disease activity and progression in MS. However, there currently are no validated clinical tests that leverage multiple blood biomarkers to track disease activity or progression in patients with MS.[18]

Development of multiplex assays can be challenging. Each protein biomarker requires specific conditions and methodologies for optimal quantification. The optimal multiplex assay should be designed so that stability and integrity of all biomarker proteins are maintained and optimized to eliminate cross-reactivity.[19] Larger scale, proteomic techniques allow higher throughput of samples and more timely readout. However, maintaining robustness, repeatability, and sensitivity is challenging, yet critical, to the validation of a multiplex biomarker panel.[20]

Analysis of multiple proteins may more accurately represent the various pathways, processes, and cell types involved in complex disease states and has the potential to deliver more personalized medicine for MS.[20-23] Single proteins may not perform well alone as a diagnostic or prognostic markers. However, as part of a multiplex assay, they may contribute to a clinically useful model when combined with other proteins and biomarkers.[21] Therefore, multiplex assay platforms have been characterized and validated for complex disease states.[19,21,22,24]

The MSDA Test is a serum-based multiplex protein biomarker assay designed to quantitatively measure disease activity using the protein levels of biomarkers present in the serum of patients with MS. Our custom multiplex assay panel was

developed using the Olink® PEA (Olink Proteomics, Uppsala, Sweden) methodology described previously (**Figure S1**).[19] Herein, we describe the comprehensive analytical characterization and validation of the MSDA Test to satisfy the criteria necessary for a fit-for-purpose validation considering the assay's intended clinical use.

2 EXPERIMENTAL SECTION

2.1 Assay development

Twenty-one biomarkers were included based on univariate and multivariate associations with clinical and radiographic MS endpoints (**Table S1**). Biological pathway modeling and protein network analysis were performed to ensure comprehensive representation of MS neurophysiology.[25]

Serum pools (*n*=4) were included on all runs during assay discovery and development. They were procured in large volumes, aliquoted, stored at −65°C, and run, in triplicate. They served as process controls to assess acceptability of future analytical runs. The SD of repeated measurements was applied to the expected concentrations. Two assay kit lots of the panel were manufactured for which critical reagents were varied to the extent possible.

2.2 Description of the two-layer stacked classifier algorithm for determination of the overall Disease Activity score

A two-layer logistic regression stacked classifier model was developed and clinically validated in a separate study that optimized the model's performance to classify serum samples based on the number of gadolinium-enhancing lesions (0 or ≥1) on an MRI administered within 60 days of blood draw.[26] The first layer of the model used the concentrations of individual, demographically corrected proteins as inputs

into four Disease Pathway models (Immunomodulation, Neuroinflammation, Myelin Biology, and Neuroaxonal Integrity). The second layer of the model used the output (eg, probabilities) of the Disease Pathway models as meta features to calculate an overall Disease Activity score (**File S1, Supporting Information**). Thresholds were established, which corresponded to low (1.0–4.0), moderate (4.5–7.0) and high (7.5–10.0) Disease Activity scores. Analytical characterization and validation of the individual biomarkers were factors used to determine inclusion of those biomarkers in the algorithm.

2.3 Incurred sample reanalysis

Incurred sample reanalysis was performed to characterize precision and robustness for the individual biomarkers and the Disease Activity and Disease Pathway scores. Forty-eight individual samples from patients with MS were repeatedly analyzed across 10 plates over ≥5 days with varied equipment, reagents, location, and personnel. Acceptability criteria for individual biomarkers was an average %CV ≤20%, and average SD at all established Disease Activity score levels of ≤1.0 units. The 48 samples broadly represented the expected range of biomarker values and Disease Activity scores in the real-world MS population.

2.4 Assay accuracy, precision, and sensitivity

Accuracy for each analyte was determined by mixing serum samples at different ratios and evaluating the percent recovery of the observed concentration relative to the expected concentration. Sample mixing enabled the accuracy assessment to be performed using endogenous protein versus a recombinant protein source. Expected concentrations were calculated by applying the targeted ratios of unmixed samples. The ratios of sample mixtures with two samples were 25%:75%, 50%:50%, and 75%:25%. The ratios of sample blends for mixtures with four samples were

25%:25%:25%:25% and 40%:10%:40%:10%. Additionally, accuracy was also evaluated for the Disease Pathway and Disease Activity algorithms by correlating observed scores with expected scores using the same sample mixtures created for the individual analyte assessments.

Intra- and inter-assay precision was measured for each analyte. The %CV was determined using serum pools enabling the assessment to be performed using endogenous protein. Serum pools were manufactured to represent patients with shorter and longer MS disease duration, those with inflammatory disease (rheumatoid arthritis), and one healthy control. Acceptability criteria for intra- and inter-assay precision was established as %CV ≤15% and ≤20%, respectively.

Sensitivity was defined as the assay's ability to accurately and precisely detect low concentrations of a given substance in biological specimens. To establish the ULOQ and LLOQ, a LOQ panel was manufactured during assay development. For each analyte, four levels were targeted near the anticipated upper limit (ULOQ 1–4) and four levels were targeted near the anticipated lower limit (LLOQ 5–8). The targeted concentrations were based on expected real-world MS patient sample distributions, the shape of the standard curve, and location of asymptotes. The LOQ panel was run in triplicate over two lots (≥5 runs per lot) and fit to the standard curve. Accuracy, defined as 80–120% recovery relative to the expected concentration and precision (inter-assay %CV ≤20%), were used to establish the acceptability criteria and determine the LLOQ and ULOQ of each analyte. Additionally, individual LOQs were assessed and established separately for each kit lot. The most conservative LOQ levels with acceptable accuracy and precision parameters for both lots were used to establish the final LLOQ and ULOQ.

Undiluted serum samples were run in the MSDA Test and as a result, no dilution factor was accounted for in the sensitivity analysis. Therefore, the LLOQ and

ULOQ define both the analytical measurement range and the reportable range of the assay. Serum samples that recovered either above the ULOQ or below the LLOQ are reported at the established LOQ concentration. MS serum samples were analyzed during the assay development process and used to establish MS reference ranges for each analyte. The linear interpolation method was used to establish the 95% interval (2.5th and 97.5th percentiles).[27] The percentile relative to these reference ranges are presented with their protein concentrations.

2.5 Assay interference

Assay interference was defined as the effect of a substance present in the sample altering the correct value of the result or the recovery of samples in the assay. Since patients with MS may be treated with a variety of drugs, potential interference of drugs was tested to determine if their presence would affect measurement of the individual protein biomarkers. Concentrations of common prescriptions, over-the-counter drugs, common MS drugs, and DMTs were spiked into serum samples (**Table S2**). Concentrations of common prescription and over-the-counter drugs were determined by Sun Diagnostics (New Gloucester, ME, USA) using a commercially available test kit. DMTs were targeted at two times C_{max} from pharmacokinetic studies, or the highest possible concentration allowable for spiking with the procured interferent stock. Finally, a universal mAb standard was tested at two concentrations (424 and 7.93 µg/mL) to cover the two times C_{max} of several mAb DMTs. Endogenous substances (hemoglobin, bilirubin, and lipids) and heterophilic antibodies (RF and HAMA) were also measured. For most interferent substances, the acceptability threshold, or median recovery, for the interference assessment was established as 80–120% relative to a corresponding spike control, except for HAMA for which percent recovery of sample mixtures was evaluated (**File S2, Supporting Information**).

2.6 Diurnal variability

Patient serum samples were collected at days 1–5 and day 12 to characterize biomarker level fluctuations. For each of the six time points per patient, the %CV and the percentage difference of the observed protein concentration relative to the average concentration at all time points were calculated.

2.7 Sample stability

In an initial experiment, stability studies for four serum samples were performed to determine the effect that storage and processing conditions can have in a clinical setting. Stability was assessed at the following four temperatures: −65°C or below (−80°C), −10°C or below (−20°C), 2–8°C (4°C), and room temperature (18–25°C) at the following time points: 4 hours (for 4°C and room temperature) and days 1, 3, 7, 14, and 28 (−20°C, 4°C, and room temperature). The results from −20°C, 4°C, and room temperature were compared with the control storage condition (−80°C). In a follow-up study, the stability of storage of 14 samples was evaluated at 4°C at days 1–3 and 7 compared with a control storage condition (−80°C) to establish the duration of time samples that can be stored at 4°C. Five freeze-thaw cycles, performed at −65°C or below, were also evaluated using four MS serum samples compared with fresh samples.

3 RESULTS AND DISCUSSION

3.1 Analytical characterization and validation

Experiments were performed between July 2020 and July 2021. Fifty-one plates were run (40 and 11 plates using the first and second lots of manufactured kits, respectively).

Based on the analytical validation and characterization of individual biomarkers described below, the 18 out of 21 biomarkers that were included in the algorithm were determined to have acceptable analytical performance. GH and COL4A1 were excluded from the algorithm based on the analytical characterization studies described below. VCAN was not incorporated into the final algorithm due to biostatistical factors unrelated to analytical performance.

3.2 Incurred sample reanalysis

All individual biomarkers were determined to have a mean %CV <20% and met established acceptability criteria (**Figure 1A**). The Disease Activity score and the four Disease Pathway scores demonstrated reproducible results throughout the range of scores (**Figures 1B–1F)**. For the Disease Activity score, the average SD across 48 samples was observed to be 0.4 score units, which is less than one interval (0.5) on the reportable scale, and as a result, met acceptability criteria. Additionally, incurred sample reanalysis showed robustness and equivalency of the assay between lots and laboratories, with the exception of COL4A1 (**Table S3**).

3.3 Assay accuracy, precision, and sensitivity

Samples for the accuracy assessment were selected from an internal MS cohort (*n*=64) to target both high and low concentrations for the individual biomarkers relative to the MS population. Twenty mixed samples from four selected samples were analyzed for each biomarker. Minimum percent recovery for each biomarker ranged from 78% to 89%; the maximum percent recovery for each biomarker ranged from 99% to 124%. The median percent recovery ranged from 91% to 100% (**Figure 2A**). Additionally, the Disease Pathway and overall Disease Activity scores were calculated for both observed and expected concentrations of the various sample

mixtures. The observed calculated scores correlated with the expected scores; R^2 ≥0.85 was established as the acceptability criteria (**Figures 2B–2F**).

Twelve replicates per serum pool were analyzed on a single plate for the intra-assay precision assessment; ≤51 values per serum pool were analyzed across 51 plates spanning 2 lots of reagent kits. The intra- and inter-assay precision satisfied the criteria for meeting the precision parameter with most analytes passing the established criteria. Of note, COL4A1 was found to have inferior inter- and intra-assay precision that ranged from 7% to 47% and 15% to 59%, respectively. Based on these findings, COL4A1 was removed from consideration for inclusion in the algorithm. MS serum samples (*N*=1645) were analyzed during the assay development and validation process and used to establish the MS reference ranges for each analyte. Sensitivity analysis demonstrated that the LLOQ and ULOQ of each analyte met the sensitivity requirements established for the assay. The maximum percentage of samples requiring imputation at any LOQ was 1.8% (for NfL at LLOQ) (**Table 1**).

3.3 Assay interference

Most of biomarker interactions with interferent combinations, such as common MS drugs, DMTs, and mAbs produced a median recovery that ranged from 80% to 120% (**Figure 3**). A lower percentage recovery was observed for two biomarkers, COL4A1 and CCL20, demonstrating a potential alteration in the presence of the sample for individual drugs. COL4A1 produced a low percent recovery for several drugs that ranged from 71% to 79%, which was likely an artifact of established assay imprecision (**Figure 3** and **Figure S2**). For CCL20, cefoxitin spiked at 660 mg/dL resulted in a median percent recovery of 77% (**Figure S2**). Additional assay interferents are shown in **Figures S2** (common drugs) and **S3** (routine endogenous interferents and heterophilic antibodies).

3.4 Diurnal variability

Diurnal variation was evaluated in eight patients over six time points (**Figure S4**). Mean and median percent differences for each biomarker and patient were observed to be within ±20%; mean and median %CV was found to be <30% for 19 of the 21 biomarkers (**Table 2**). Of note, there were some individual samples that were outside of the acceptable range (±30%; data not shown). In addition, mean and median diurnal variability ≥30% was observed for COL4A1, which may have been due to the imprecision of the assay to detect this biomarker. GH was also found to be more variable compared with the other biomarkers, which is not surprising, as GH has been previously reported to have a high degree of ultradian and diurnal variability.[28] For this reason, GH was removed from consideration for inclusion in the algorithm.

3.5 Sample stability

In the initial stability study, all biomarkers were stable for up to 1 day at room temperature and at 4°C, and 28 days at −20°C. For those samples stored at room temperature, CXCL13, IL-12β, and TNFSF13B decreased beyond −20% at 3 days. During a follow-up study, all biomarkers were found to meet acceptability criteria when stored at 4°C, and consistent with the initial study as well as the control condition (−80°C) at follow-up (**Table S4**). In a study to examine the stability of samples after freeze-thaw, most biomarkers met acceptability criteria when compared with fresh sample. Of note, GFAP concentrations decreased beyond −20% for freeze-thaw cycles 4 and 5 (**Table S5**). Finally, score level analysis showed that test conditions were within 3 SDs (±1.5 score difference) from the control conditions during the initial study (**Table S6**) and at follow-up (**Table S7**). From these findings, we showed that biomarker levels were found to be most affected above certain thresholds (room temperature for 24 h, 4°C for 7 days, −20°C for 28 days, and three

freeze-thaws). These data can be used to establish allowable sample handling and storage conditions.

4 CONCLUSION

The accuracy, precision, and reproducibility of a biomarker assay are critical to its utility as a diagnostic and prognostic tool in the management of complex neurodegenerative disorders such as MS. Additionally, such an assay should be insensitive to external factors such as assay interferents and sample collection, processing, and storage. The Clinical and Laboratory Standards Institute and the United States Food and Drug Administration issued guidance on the development and validation of assays for the detection of serum-based biomarkers.[27,29,30] Parameters such as accuracy, precision, recovery, sensitivity and specificity, quality control, and sample stability need to be optimized for the assay to be properly validated.[27,29,30] Results from our analytical validation experiments to characterize the MSDA Test support that the assay is accurate, precise, sensitive, specific, and robust at determining individual biomarker levels and algorithmic scores, regardless of assay interferents, and validated in terms of sample stability. Our findings of high accuracy and precision for the MSDA Test assay align with those of other validation studies of multiplex assays utilizing the same,[31,32] as well as alternative[21,22] platforms.

PEA demonstrated high sensitivity, specificity, reproducibility, and repeatability with low intra- and inter-assay variability, which has allowed for large-scale, high throughput screening of up to 92 proteins in 96 samples simultaneously, with low sample consumption and cost.[19] This platform detected novel protein biomarkers and biomarker combinations for many complex disease states, such as

cardiovascular disease,[33-37] cancer,[32,38-40] Alzheimer's disease,[41] and inflammatory disease states, such as atopic dermatitis and lupus,[42,43] as well as in aging research.[44] For the MSDA Test, we demonstrated that a focused panel of MS biomarkers can be developed and optimized on the PEA platform with absolute quantitation of the proteins to support a fit-for-purpose analytical validation, thereby enabling clinical use of the assay.

Thus far, there are no validated clinical tests that leverage multiple blood biomarkers to track disease activity or progression in patients with MS. This is critical for a disease such as MS, which has a complicated clinical course varying from mild, self-limiting to severe.[1] Although MS disease prognosis is primarily based on clinical evidence, such as relapse rate and disability progression, and diagnostic tests (eg, brain MRI or the presence of oligoclonal immunoglobulin G bands in the CSF),[14] neither can consistently and accurately predict disease course, activity, or prognosis.[13] Given the emphasis on early diagnosis and the efficacy of therapies to treat early stages of the relapsing/remitting form of the disease,[4,5] validation of a biomarker panel remains an unmet need in clinical practice, and use of this biomarker tool should provide diagnostic and prognostic value for the treatment of MS. This study demonstrated identification of biomarkers for this complex disease using the PEA platform. With further clinical validation, this assay can be used to track disease activity and progression of MS, allowing a more personalized approach to MS treatment.

A limitation of using a multiplex assay is that the conditions established for one biomarker are not always uniform across the full panel of biomarkers. Our findings show that the MSDA Test was optimized for assessment of 18 out of the 21 included biomarkers and the analytical validation paradigm that we described demonstrates a

high level of accuracy, sensitivity, and precision with minimal cross-reactivity and interference by substances commonly seen in patients with MS.

This study serves as a critical first step in the validation of a multivariate proteomic serum-based assay. The next step in the validation of the MSDA Test is clinical validation, which will support and confirm the association between the serum-based MSDA Test and clinical and radiographic MS endpoints. Upon completion of clinical validation of the assay, the final Disease Activity and Disease Pathway algorithms will use the ensemble of validated proteins to expand the use of the assay by evaluating biomarker correlations with endpoints associated with additional MS disease assessments, selection of therapy, and differential diagnosis of patients with MS. Upon successful clinical validation, this MSDA Test will be a quantitative, minimally invasive, and scalable tool to improve disease management for patients with MS and their physicians.

ACKNOWLEDGMENTS

The authors wish to thank the following team members from Olink Proteomics (Uppsala, Sweden) who were involved in the development of the Multiple Sclerosis Disease Activity Test assay (Erika Assarsson, Sandra Ohlsson, Martin Lundberg, Jessica Bergman, and Niklas Nordberg). All authors contributed to and approved the manuscript for submission. Writing and editorial assistance were provided by Jennifer L. Venzie, PhD, of The Lockwood Group (Stamford, CT), and were funded by Octave Bioscience, Inc.

CONFLICT OF INTEREST STATEMENT

F. Qureshi, W. Hu, L. Loh, H. Patel, M. DeGuzman, F. Rubio da Costa, V. Gehman, and F. Zhang are employees of Octave Bioscience, Inc. M. Becich was an employee of Octave Bioscience, Inc. at the time the study was completed. J. Foley has received research support from Biogen, Novartis, Adamas, Octave Bioscience, Inc., Genentech, and Mallinckrodt, received speakers' honoraria and acted as a consultant for EMD Serono, Genzyme, Novartis, Biogen, and Genentech, and has equity interest in Octave Bioscience Inc., and is the founder of InterPro Bioscience. T. Chitnis has received compensation for consulting from Biogen, Novartis Pharmaceuticals, Roche Genentech, and Sanofi Genzyme, and received research support from the National Institutes of Health, National MS Society, US Department of Defense, EMD Serono, I-Mab Biopharma, Mallinckrodt ARD, Novartis Pharmaceuticals, Octave Bioscience, Inc., Roche Genentech, and Tiziana Life Sciences.

FUNDING

The study was funded by Octave Bioscience, Inc. and in part by the U.S. Department of Defense (W81XWH2110633 to T Chitnis).

DATA AVAILABILITY STATEMENT

Access to data can be provided after a research proposal is submitted to the corresponding author and a data sharing agreement is in place.

REFERENCES

[1] Compston, A., Coles, A. (2008). Multiple sclerosis. *Lancet*, *372*, 1502-1517.

[2] Weiner, H. L. (2004). Multiple sclerosis is an inflammatory T-cell–mediated autoimmune disease. *Arch Neurol*, *61*, 1613-1615.

[3] Miller, D. H., Chard, D. T., Ciccarelli, O. (2012). Clinically isolated syndromes. *Lancet Neurol*, *11*, 157-169.

[4] Giovannoni, G., Butzkueven, H., Dhib-jalbut, S., Hobart, J., Kobelt, G., Pepper, G., . . . Vollmer, T. (2016). Brain health: time matters in multiple sclerosis. *Mult Scler Relat Disord*, *9*, S5-S48.

[5] Macaron, G., Ontaneda, D. (2019). Diagnosis and management of progressive multiple sclerosis. *Biomedicines*, *7*, 56.

[6] Miller, J. R. (2004). The importance of early diagnosis of multiple sclerosis. *J Manag Care Pharm*, *10*, S4-S11.

[7] Goldschmidt, T., Antel, J., König, F. B., Brück, W., Kuhlmann, T. (2009). Remyelination capacity of the MS brain decreases with disease chronicity. *Neurology*, *72*, 1914-1921.

[8] Cunniffe, N., Coles, A. (2021). Promoting remyelination in multiple sclerosis. *J Neurol*, *268*, 30-44.

[9] McDonald, W. I., Compston, A., Edan, G., Goodkin, D., Hartung, H.-P., Lublin, F. D., . . . Wolinsky, J. S. (2001). Recommended diagnostic criteria for multiple sclerosis: guidelines from the international panel on the diagnosis of multiple sclerosis. *Ann Neurol*, *50*, 121-127.

[10] Thompson, A. J., Banwell, B. L., Barkhof, F., Carroll, W. M., Coetzee, T., Comi, G., . . . Cohen, J. A. (2018). Diagnosis of multiple sclerosis: 2017 revisions of the McDonald criteria. *Lancet Neurol*, *17*, 162-173.

[11] Tintoré, M., Rovira, A., Río, J., Nos, C., Grivé, E., Téllez, N., . . . Montalban, X. (2006). Baseline MRI predicts future attacks and disability in clinically isolated syndromes. *Neurology*, *67*, 968-972.

[12] Teixeira, M., Seabra, M., Carvalho, L., Sequeira, L., Abreu, P., Mendonça, T., . . . Guimarães, J. (2020). Clinically isolated syndrome, oligoclonal bands and multiple sclerosis. *Clin Exp Neuroimmunol*, *11*, 33-39.

[13] Jafari, A., Babajani, A., Rezaei-Tavirani, M. (2021). Multiple sclerosis biomarker discoveries by proteomics and metabolomics approaches. *Biomark Insights*, *16*, 11772719211013352.

[14] Ziemssen, T., Akgün, K., Brück, W. (2019). Molecular biomarkers in multiple sclerosis. *J Neuroinflammation*, *16*, 272.

[15] Gul, M., Jafari, A. A., Shah, M., Mirmoeeni, S., Haider, S. U., Moinuddin, S., Chaudhry, A. (2020). Molecular biomarkers in multiple sclerosis and its related disorders: a critical review. *Int J Mol Sci*, *21*, 6020.

[16] Hühmer, A. F., Biringer, R. G., Amato, H., Fonteh, A. N., Harrington, M. G. (2006). Protein analysis in human cerebrospinal fluid: physiological aspects, current progress and future challenges. *Dis Markers*, *22*, 3-26.

[17] Zhang, J. (2007). Proteomics of human cerebrospinal fluid – the good, the bad, and the ugly. *Proteomics Clin Appl*, *1*, 805-819.

[18] Lublin, F. D., Reingold, S. C., Cohen, J. A., Cutter, G. R., Sørensen, P. S., Thompson, A. J., . . . Polman, C. H. (2014). Defining the clinical course of multiple sclerosis: the 2013 revisions. *Neurology*, *83*, 278-286.

[19] Assarsson, E., Lundberg, M., Holmquist, G., Björkesten, J., Bucht Thorsen, S., Ekman, D., . . . Fredriksson, S. (2014). Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS One*, *9*, e95192.

[20] Schubert, O. T., Röst, H. L., Collins, B. C., Rosenberger, G., Aebersold, R. (2017). Quantitative proteomics: challenges and opportunities in basic and applied research. *Nat Protoc*, *12*, 1289-1294.

[21] Dillon, R., Croner, L. J., Bucci, J., Kairs, S. N., You, J., Beasley, S., . . . Wilcox, B. (2018). Analytical validation of a novel multiplex test for detection of advanced adenoma and colorectal cancer in symptomatic patients. *J Pharm Biomed Anal*, *154*, 85-94.

[22] Eastman, P. S., Manning, W. C., Qureshi, F., Haney, D., Cavet, G., Alexander, C., Hesterberg, L. K. (2012). Characterization of a multiplex, 12-biomarker test for rheumatoid arthritis. *J Pharm Biomed Anal*, *70*, 415-424.

[23] Chitnis, T., Prat, A. (2020). A roadmap to precision medicine for multiple sclerosis. *Multiple Sclerosis Journal*, *26*, 522-532.

[24] Bhawal, R., Oberg, A. L., Zhang, S., Kohli, M. (2020). Challenges and opportunities in clinical applications of blood-based proteomics in cancer. *Cancers (Basel)*, *12*, 2428.

[25] Chitnis, T., Becich, M., Bove, R., Cree, B. A. C., Gehman, V., Gomez, R., . . . Baranzini, S. E. Development of a custom multivariate proteomic serum based assay for association with radiographic and clinical endpoints in MS. Presented at: Americas Committee for Treatment and Research in Multiple Sclerosis and European Committee for Treatment and Research in Multiple Sclerosis; September 11-13, 2020; Virtual.

[26] Chitnis, T., Foley, J., Ionete, C., El Ayoubi, N., Saxena, S., Gaitan-Walsh, P., . . . Khoury, S. J. Clinical validation study results of a multivariate proteomic serum based assay for disease activity assessments in multiple sclerosis. Presented at: 37th Congress of the European Committee for Treatment and Research in Multiple Sclerosis (ECTRIMS); October 13-15, 2021; Virtual.

[27] Clinical and Laboratory Standards Institute. (2018). EP283CE. Defining, establishing, and verifying reference intervals in the clinical laboratory. 3rd edition.

[28] Hancox, T. P. M., Skene, D. J., Dallmann, R., Dunn, W. B. (2021). Tick-tock consider the clock: the influence of circadian and external cycles on time of day variation in the human metabolome—a review. *Metabolites*, *11*, 328.

[29] United States Food and Drug Administration. Bioanalytical method validation: guidance for industry. Accessed September 13, 2021. https://www.fda.gov/media/70858/download

[30] Clinical and Laboratory Standards Institute. (2016). EP07. Interference testing in clinical chemistry. 3rd edition.

[31] Xu, Y. W., Peng, Y. H., Chen, B., Wu, Z. Y., Wu, J. Y., Shen, J. H., . . . Xu, L. Y. (2014). Autoantibodies as potential biomarkers for the early detection of esophageal squamous cell carcinoma. *Am J Gastroenterol*, *109*, 36-45.

[32] Yang, X., Suo, C., Zhang, T., Yin, X., Man, J., Yuan, Z., . . . Ye, W. (2021). Targeted proteomics-derived biomarker profile develops a multi-protein classifier in liquid biopsies for early detection of esophageal squamous cell carcinoma from a population-based case-control study. *Biomark Res*, *9*, 12.

[33] Hijazi, Z., Wallentin, L., Lindbäck, J., Alexander, J. H., Connolly, S. J., Eikelboom, J. W., . . . Siegbahn, A. (2020). Screening of multiple biomarkers associated with ischemic stroke in atrial fibrillation. *J Am Heart Assoc*, *9*, e018984.

[34] Sanders-van Wijk, S., Tromp, J., Beussink-Nelson, L., Hage, C., Svedlund, S., Saraste, A., . . . Shah, S. J. (2020). Proteomic evaluation of the comorbidity-inflammation paradigm in heart failure with preserved ejection fraction: results from the PROMIS-HFpEF study. *Circulation*, *142*, 2029-2044.

[35] Pol, T., Hijazi, Z., Lindbäck, J., Oldgren, J., Alexander, J. H., Connolly, S. J., . . . Wallentin, L. (2021). Using multimarker screening to identify biomarkers associated with cardiovascular death in patients with atrial fibrillation. *Cardiovasc Res*, cvab262.

[36] Feldreich, T., Nowak, C., Fall, T., Carlsson, A. C., Carrero, J. J., Ripsweden, J., . . . Ärnlöv, J. (2019). Circulating proteins as predictors of cardiovascular mortality in end-stage renal disease. *J Nephrol*, *32*, 111-119.

[37] Nowak, C., Carlsson, A. C., Östgren, C. J., Nyström, F. H., Alam, M., Feldreich, T., . . . Ärnlöv, J. (2018). Multiplex proteomics for prediction of major cardiovascular events in type 2 diabetes. *Diabetologia*, *61*, 1748-1757.

[38] Leandersson, P., Åkesson, A., Hedenfalk, I., Malander, S., Borgfeldt, C. (2020). A multiplex biomarker assay improves the diagnostic performance of HE4 and CA125 in ovarian tumor patients. *PLoS One*, *15*, e0240418.

[39] Enroth, S., Berggrund, M., Lycke, M., Broberg, J., Lundberg, M., Assarsson, E., . . . Gyllensten, U. (2019). High throughput proteomics identifies a high-accuracy 11 plasma protein biomarker signature for ovarian cancer. *Commun Biol*, *2*, 221.

[40] Berggrund, M., Enroth, S., Lundberg, M., Assarsson, E., Stålberg, K., Lindquist, D., . . . Gyllensten, U. (2019). Identification of candidate plasma protein biomarkers for cervical cancer using the multiplex proximity extension assay. *Mol Cell Proteomics*, *18*, 735-743.

[41] Whelan, C. D., Mattsson, N., Nagle, M. W., Vijayaraghavan, S., Hyde, C., Janelidze, S., . . . Hansson, O. (2019). Multiplex proteomics identifies novel CSF and plasma biomarkers of early Alzheimer's disease. *Acta Neuropathol Commun*, *7*, 169.

[42] Petrackova, A., Smrzova, A., Gajdos, P., Schubertova, M., Schneiderova, P., Kromer, P., . . . Kriegova, E. (2017). Serum protein pattern associated with organ damage and lupus nephritis in systemic lupus erythematosus revealed by PEA immunoassay. *Clin Proteomics*, *14*, 32.

[43] Brunner, P. M., He, H., Pavel, A. B., Czarnowicki, T., Lefferdink, R., Erickson, T., . . . Paller, A. S. (2019). The blood proteomic signature of early-onset pediatric atopic dermatitis shows systemic inflammation and is distinct from adult long-standing disease. *J Am Acad Dermatol*, *81*, 510-519.

[44] Moaddel, R., Ubaida-Mohien, C., Tanaka, T., Lyashkov, A., Basisty, N., Schilling, B., . . . Ferrucci, L. (2021). Proteomics in aging research: a roadmap to clinical, translational research. *Aging Cell*, *20*, e13325.

FIGURE LEGENDS

FIGURE 1. Incurred sample reanalysis results for (A) Individual biomarkers and (B) Overall Disease Activity score, (C) Immunomodulation, (D) Neuroinflammation, (E) Myelin Biology, and (F) Neuroaxonal Integrity pathway scores in the MSDA Test.

FIGURE 2. Accuracy of the MSDA Test to detect (A) Individual biomarkers and (B) Overall Disease Activity score, (C) Immunomodulation, (D) Neuroinflammation, (E) Myelin Biology, and (F) Neuroaxonal Integrity pathway scores.

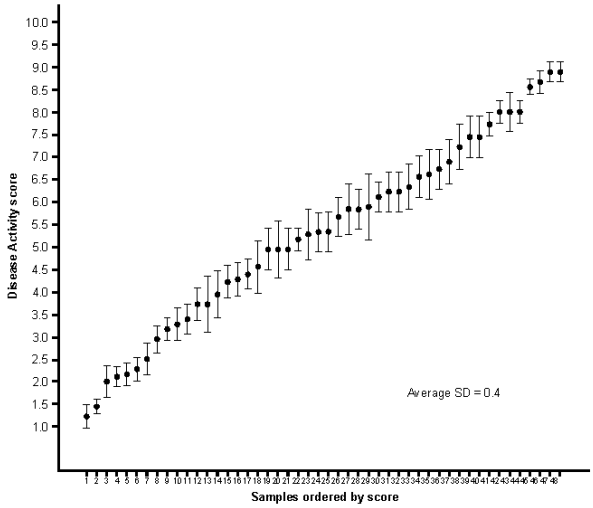
FIGURE 3. Assay interference for common MS drugs, DMTs, and the high concentration of universal mAb surrogates in the MSDA Test.

FIGURE 1

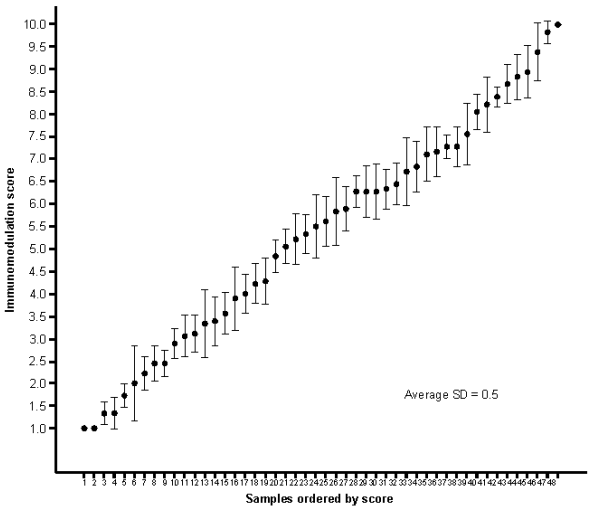
A. **Individual biomarkers**

Analyte	Mean (% CV)
APLP1	13
CCL20	13
CD6	11
CDCP1	12
CNTN2	11
COL4A1	18
CXCL13	10
CXCL9	12
FLRT2	10
GFAP	19
GH	12
IL-12β	11
MOG	9
Nfl	14
OPG	13
OPN	11
PRTG	8
SERPINA9	11
TNFRSF10A	12
TNFSF13B	11
VCAN	9

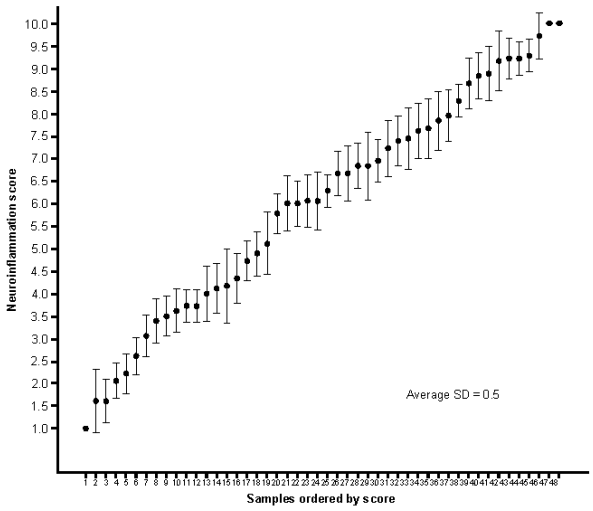
B. **Disease Activity score**



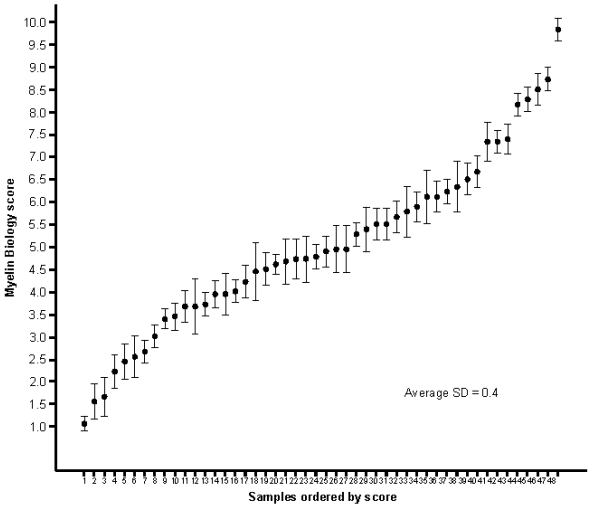
C. **Immunodulation score**



D. **Neuroinflammation score**



E. **Myelin Biology score**



F. **Neuroaxonal Integrity score**

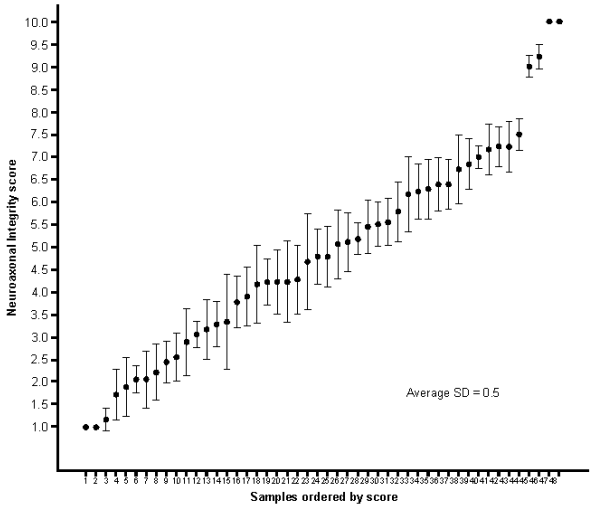


FIGURE 2

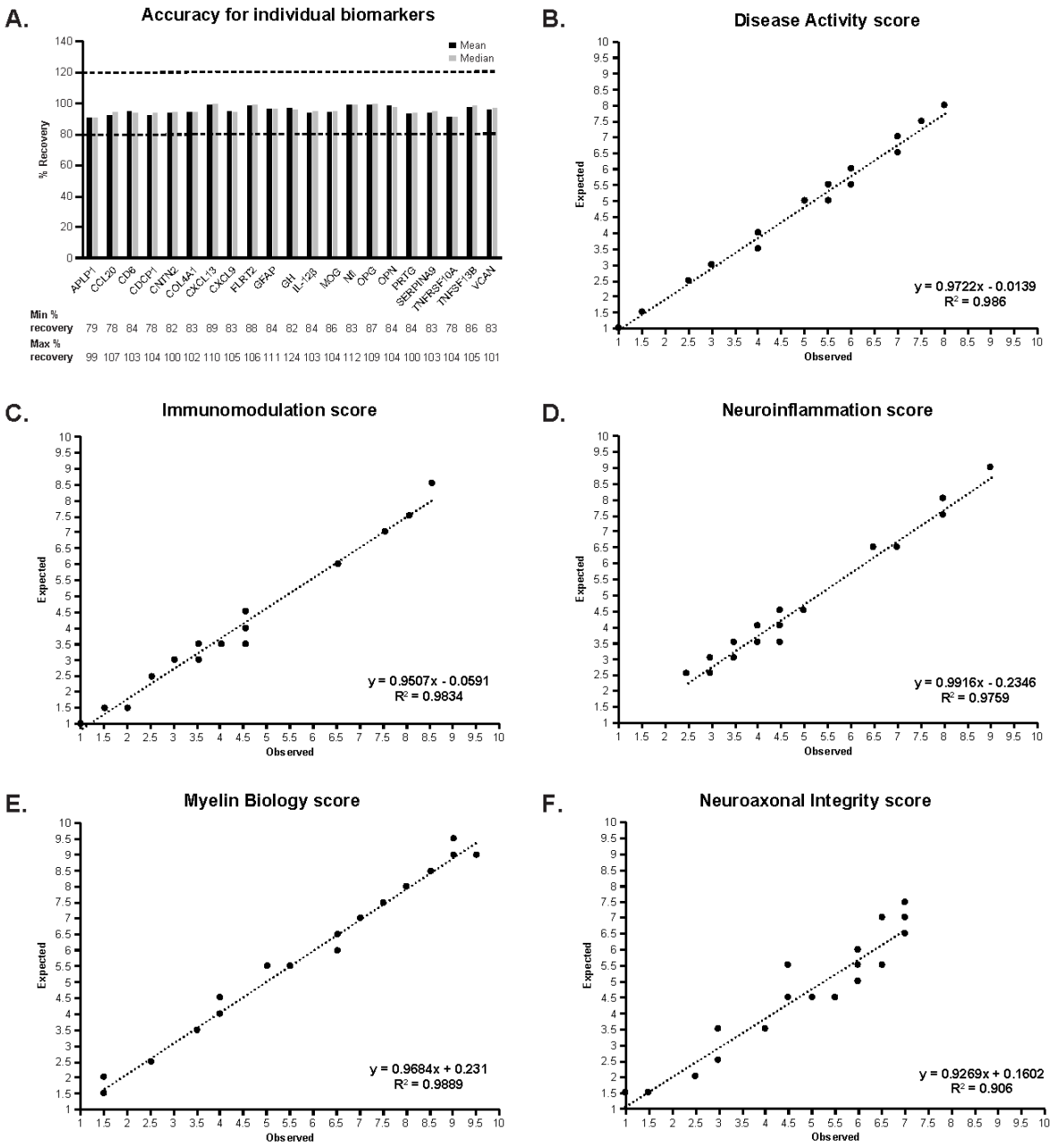

A.
Accuracy for individual biomarkers
Mean
Median
% Recovery
APLP1 COL20 CD6 CIDCP1 CNTN2 COL4A1 CXCL13 CXCL9 FLRT2 GFAP GH IL-12B MOG NM OR2 OPN PRTG SERPINA9 TNFRSF10A TNFSF13B VCAN
Min % recovery 79 78 84 78 82 83 89 83 88 84 82 84 86 83 87 84 84 83 78 86 83
Max % recovery 99 107 103 104 100 102 110 105 106 111 124 103 104 112 109 104 100 103 104 105 101
B.
Disease Activity score
Expected
Observed
y = 0.9722x - 0.0139
R² = 0.986
C.
Immunomodulation score
Expected
Observed
y = 0.9507x - 0.0591
R² = 0.9834
D.
Neuroinflammation score
Expected
Observed
y = 0.9916x - 0.2346
R² = 0.9759
E.
Myelin Biology score
Expected
Observed
y = 0.9684x + 0.231
R² = 0.9889
F.
Neuroaxonal Integrity score
Expected
Observed
y = 0.9269x + 0.1602
R² = 0.906

FIGURE 3

% Recovery
0
20
40
60
80
100
120
140
Interferon β-1A
Dimethyl fumarate
Fingolimod
Teriflunomide
Mitoxantrone
Glatiramer acetate
Rituximab
Methylprednisolone
Cladribine
Vitamin D
Universal mAb (high)
APLPL1
CCL20
CD6
CDCP1
CNTN2
COL4A1
CXCL13
CXCL9
FLRT2
GFAP
GH
IL-12β
MOG
Nfl
OPG
OPN
PRTG
SERPINA9
TNFRSF10A
TNFSF13B
VCAN

TABLES

TABLE 1 Intra- and inter-assay precision, sensitivity, and reference ranges for biomarkers in the MSDA Test

Analytes	Precision												Sensitivity and reference ranges					
	Shorter MS duration pool[a]			Longer MS duration pool[b]			RA pool			Healthy control pool								
	Intra %CV	Inter %CV	Conc (pg/mL)	Intra %CV	Inter %CV	Conc (pg/mL)	Intra %CV	Inter %CV	Conc (pg/mL)	Intra %CV	Inter %CV	Conc (pg/mL)	LLOQ (pg/mL)	ULOQ (pg/mL)	Low MS range (pg/mL)[c]	High MS range (pg/mL)[d]	Samples imputed at LLOQ, % (*N*=1645)	Samples imputed at ULOQ, % (*N*=1645)
APLP1	9	9	10,296	9	8	11,560	4	8	11,868	7	9	11,868	2323.78	142,798.49	5500	22,000	0	0
CCL20	6	9	6.9	8	7	9.2	5	9	13.7	7	9	11.8	0.92	383.49	2.1	52	0	1 (<0.1)
CD6	6	8	89	8	8	108	5	7	137	8	9	112	4.62	3318.60	46	250	0	0
CDCP1	8	9	78	8	9	125	4	9	208	7	10	72	24.22	6795.04	28	230	23 (1.4)	1 (<0.1)
CNTN2	6	7	1120	7	7	1643	4	6	1554	8	8	1256	44.46	12,373.51	650	3300	0	0
COL4A1	7	15	1104	19	20	1334	8	59	1601	47	27	1387	30.65	4573.38	520	3600	0	23 (1.4)
CXCL13	6	8	52.8	8	7	42.9	7	8	65.3	7	9	46.8	1.91	1112.70	22	190	0	0
CXCL9	6	11	31.0	8	10	62.6	5	11	112.3	7	11	27.5	1.89	1832.22	17	250	0	0
FLRT2	7	8	103	9	9	110	5	8	139	8	9	116	35.67	10,107.17	63	180	1 (<0.1)	0
GFAP	7	18	70	10	16	126	8	15	148	9	18	77	12.46	19,582.88	24	220	16 (1.0)	0
GH	7	9	823	8	7	595	5	8	1010	7	9	366	9.63	18,413.83	17	9500	7 (0.4)	9 (0.5)
IL-12β	7	9	109	9	7	122	6	8	118	8	9	71	0.56	3044.33	28	280	0	0
MOG	4	6	21.9	7	6	22.8	5	7	26.0	6	8	17.8	1.75	577.42	12	47	0	0
NfL	10	11	7.6	13	9	15.6	8	8	20.6	11	12	6.5	3.31	599.18	3.5	42	29 (1.8)	0
OPG	6	11	699	9	11	806	6	10	1022	7	12	602	14.58	62,385.21	410	1400	0	0
OPN	6	10	15,733	7	10	15,415	6	12	17,470	7	13	10,450	572.50	157,267.29	9500	39,000	0	0
PRTG	7	6	94	8	7	107	5	6	108	7	7	103	3.90	5920.73	71	180	2 (0.1)	0
SERPINA9	11	8	45.1	11	8	37.9	5	7	60.0	6	14	50.0	5.12	9286.67	12	160	1 (<0.1)	0
TNFRSF10A	9	9	5.1	9	9	5.5	5	8	7.6	9	9	4.9	0.48	1027.28	2.8	9.7	0	0
TNFSF13B	7	10	4075	11	11	4019	5	11	4204	7	13	3003	660.29	130,682.08	2300	10,000	0	0
VCAN	7	8	316	7	7	337	4	7	448	5	8	310	8.54	14,673.95	230	600	0	0

Green shading: intra- (%CV ≤15%) or inter-assay (%CV <20%) are within the acceptability range for the assays. Red shading: intra- (%CV >15%) or inter-assay (%CV ≥20%) precision values are outside the acceptability range for the assays.

[a]Average age of patients with shorter MS duration was 36 (range, 27–43) years.

[b]Average age of patients with longer MS duration was 52 (range, 45–62) years.

[c]Low MS range was defined as the 2.5th percentile.

[d]High MS range was defined as the 97.5th percentile.

TABLE 2 Diurnal variability of eight samples in the MSDA Test across six time points (days 1, 2, 3, 4, 5, and 12)

Analyte	Mean %CV	Median %CV
APLP1	14	15
CCL20	25	21
CD6	10	7
CDCP1	12	12
CNTN2	11	10
COL4A1	44	39
CXCL13	18	11
CXCL9	13	11
FLRT2	10	8
GFAP	13	12
GH	78	79
IL-12β	10	9
MOG	12	11
NfL	17	17
OPG	11	10
OPN	10	8
PRTG	8	6
SERPINA9	12	12
TNFRSF10A	10	12
TNFSF13B	11	8
VCAN	10	7

Green shading: %CV ≤30%. Red shading: %CV >30%.