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 Chitnis3

1Octave Bioscience, Inc., Menlo Park, CA, USA
2Rocky Mountain Multiple Sclerosis Clinic, Salt Lake City, UT, USA
3Brigham 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; Cmax, 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

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
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 $C_{\text{max}}$ 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 $C_{\text{max}}$ 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; $R^2 \geq 0.85$ 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; $\leq 51$ 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


Qureshi et al. Biomarker Analytical Validation Manuscript


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

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Mean (LOD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACET</td>
<td>12</td>
</tr>
<tr>
<td>CCL20</td>
<td>13</td>
</tr>
<tr>
<td>CD6</td>
<td>11</td>
</tr>
<tr>
<td>CD21</td>
<td>10</td>
</tr>
<tr>
<td>CNOT1Q2</td>
<td>11</td>
</tr>
<tr>
<td>CCL15</td>
<td>10</td>
</tr>
<tr>
<td>C1S</td>
<td>10</td>
</tr>
<tr>
<td>CXCL13</td>
<td>12</td>
</tr>
<tr>
<td>FLTR2</td>
<td>10</td>
</tr>
<tr>
<td>GLP1R</td>
<td>10</td>
</tr>
<tr>
<td>GM1</td>
<td>12</td>
</tr>
<tr>
<td>G4.1Q11</td>
<td>11</td>
</tr>
<tr>
<td>APOJ</td>
<td>9</td>
</tr>
<tr>
<td>APOE</td>
<td>14</td>
</tr>
<tr>
<td>CRP</td>
<td>13</td>
</tr>
<tr>
<td>CRP</td>
<td>11</td>
</tr>
<tr>
<td>FANCA</td>
<td>9</td>
</tr>
<tr>
<td>FENITL</td>
<td>8</td>
</tr>
<tr>
<td>STEAP1MA2</td>
<td>11</td>
</tr>
<tr>
<td>TNFRSF11A</td>
<td>12</td>
</tr>
<tr>
<td>TAF15Y1</td>
<td>11</td>
</tr>
<tr>
<td>TAF15Y1B</td>
<td>11</td>
</tr>
</tbody>
</table>

B. Disease Activity score

Average SD = 0.4

Sample ordered by score

C. Immunodulation score

Average SD = 0.5

Sample ordered by score

D. Neuroinflammation score

Average SD = 0.5

Sample ordered by score

E. Myelin Biology score

Average SD = 0.4

Sample ordered by score

F. Neuroaxonal Integrity score

Average SD = 0.6

Sample ordered by score
FIGURE 2

A. Accuracy for Individual biomarkers

B. Disease Activity score

C. Immunomodulation score

D. Neuroinflammation score

E. Myelin Biology score

F. Neuroaxonal Integrity score
**TABLE 1** Intra- and inter-assay precision, sensitivity, and reference ranges for biomarkers in the MSDA Test

<table>
<thead>
<tr>
<th>Analytes</th>
<th>Shorter MS duration pool</th>
<th>Longer MS duration pool</th>
<th>RA pool</th>
<th>Healthy control pool</th>
<th>LLOQ (pg/mL)</th>
<th>ULOQ (pg/mL)</th>
<th>Low MS range (pg/mL)</th>
<th>High MS range (pg/mL)</th>
<th>Samples imputed at LLOQ, % (N=1645)</th>
<th>Samples imputed at ULOQ, % (N=1645)</th>
</tr>
</thead>
<tbody>
<tr>
<td>APLP1</td>
<td>9</td>
<td>9</td>
<td>10,296</td>
<td>9</td>
<td>8</td>
<td>11,500</td>
<td>11,868</td>
<td>7</td>
<td>11,868</td>
<td>2323.78</td>
</tr>
<tr>
<td>CCL20</td>
<td>6</td>
<td>9</td>
<td>6.9</td>
<td>8</td>
<td>8</td>
<td>9.2</td>
<td>13.7</td>
<td>7</td>
<td>9</td>
<td>0.92</td>
</tr>
<tr>
<td>CD6</td>
<td>6</td>
<td>8</td>
<td>8.9</td>
<td>8</td>
<td>8</td>
<td>108</td>
<td>157</td>
<td>4</td>
<td>8</td>
<td>4.62</td>
</tr>
<tr>
<td>CDCP1</td>
<td>8</td>
<td>9</td>
<td>78</td>
<td>8</td>
<td>9</td>
<td>125</td>
<td>208</td>
<td>7</td>
<td>10</td>
<td>24.22</td>
</tr>
<tr>
<td>CNTN2</td>
<td>6</td>
<td>7</td>
<td>1120</td>
<td>7</td>
<td>7</td>
<td>1643</td>
<td>1554</td>
<td>8</td>
<td>8</td>
<td>44.46</td>
</tr>
<tr>
<td>COL4A1</td>
<td>7</td>
<td>15</td>
<td>1104</td>
<td>19</td>
<td>20</td>
<td>1343</td>
<td>1601</td>
<td>47</td>
<td>27</td>
<td>30.65</td>
</tr>
<tr>
<td>CXCL13</td>
<td>6</td>
<td>8</td>
<td>52.8</td>
<td>8</td>
<td>7</td>
<td>42.9</td>
<td>65.3</td>
<td>7</td>
<td>9</td>
<td>1.91</td>
</tr>
<tr>
<td>CXCL9</td>
<td>6</td>
<td>11</td>
<td>31.0</td>
<td>8</td>
<td>10</td>
<td>62.6</td>
<td>112.3</td>
<td>7</td>
<td>11</td>
<td>1.89</td>
</tr>
<tr>
<td>FLRT2</td>
<td>7</td>
<td>8</td>
<td>103</td>
<td>9</td>
<td>9</td>
<td>110</td>
<td>139</td>
<td>8</td>
<td>9</td>
<td>35.67</td>
</tr>
<tr>
<td>GFAP</td>
<td>7</td>
<td>18</td>
<td>70</td>
<td>10</td>
<td>16</td>
<td>126</td>
<td>148</td>
<td>9</td>
<td>18</td>
<td>12.46</td>
</tr>
<tr>
<td>GH</td>
<td>7</td>
<td>9</td>
<td>823</td>
<td>8</td>
<td>7</td>
<td>595</td>
<td>1010</td>
<td>7</td>
<td>9</td>
<td>9.63</td>
</tr>
<tr>
<td>IL-12β</td>
<td>7</td>
<td>9</td>
<td>109</td>
<td>9</td>
<td>7</td>
<td>122</td>
<td>118</td>
<td>8</td>
<td>9</td>
<td>0.56</td>
</tr>
<tr>
<td>MOG</td>
<td>4</td>
<td>6</td>
<td>21.9</td>
<td>7</td>
<td>6</td>
<td>22.8</td>
<td>26.0</td>
<td>6</td>
<td>8</td>
<td>1.75</td>
</tr>
<tr>
<td>NFL</td>
<td>10</td>
<td>11</td>
<td>7.6</td>
<td>13</td>
<td>9</td>
<td>15.6</td>
<td>20.6</td>
<td>11</td>
<td>12</td>
<td>3.31</td>
</tr>
<tr>
<td>OPG</td>
<td>6</td>
<td>11</td>
<td>699</td>
<td>9</td>
<td>11</td>
<td>806</td>
<td>1022</td>
<td>7</td>
<td>12</td>
<td>14.58</td>
</tr>
<tr>
<td>OPN</td>
<td>6</td>
<td>10</td>
<td>15,733</td>
<td>7</td>
<td>10</td>
<td>15,415</td>
<td>17,470</td>
<td>7</td>
<td>13</td>
<td>572.50</td>
</tr>
<tr>
<td>PRTG</td>
<td>7</td>
<td>6</td>
<td>94</td>
<td>8</td>
<td>7</td>
<td>107</td>
<td>108</td>
<td>7</td>
<td>7</td>
<td>3.90</td>
</tr>
<tr>
<td>SERPINA9</td>
<td>11</td>
<td>8</td>
<td>45.1</td>
<td>11</td>
<td>8</td>
<td>37.9</td>
<td>60.0</td>
<td>6</td>
<td>14</td>
<td>5.12</td>
</tr>
<tr>
<td>TNFRSF10A</td>
<td>9</td>
<td>9</td>
<td>5.1</td>
<td>9</td>
<td>9</td>
<td>5.5</td>
<td>7.6</td>
<td>9</td>
<td>9</td>
<td>0.48</td>
</tr>
<tr>
<td>TNFSF13B</td>
<td>7</td>
<td>10</td>
<td>4075</td>
<td>11</td>
<td>11</td>
<td>4019</td>
<td>4204</td>
<td>7</td>
<td>13</td>
<td>660.29</td>
</tr>
<tr>
<td>VCAN</td>
<td>7</td>
<td>8</td>
<td>316</td>
<td>7</td>
<td>7</td>
<td>337</td>
<td>448</td>
<td>5</td>
<td>8</td>
<td>8.54</td>
</tr>
</tbody>
</table>

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

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

Low MS range was defined as the 2.5th percentile.  

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)

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

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