Analytical Validation of a Multivariate Proteomic Serum Based Assay for Disease Activity Assessments in Multiple Sclerosis

W. Hu\textsuperscript{1}, L. Loh\textsuperscript{1}, H. Patel\textsuperscript{1}, M. DeGuzman\textsuperscript{1}, M. Becich\textsuperscript{1}, F. Rubio da Costa\textsuperscript{1}, V. Gehman\textsuperscript{1}, E. Assarsson\textsuperscript{2}, S.Ohlsson\textsuperscript{2}, M. Lundberg\textsuperscript{2}, J. Bergman\textsuperscript{2}, N. Nordberg\textsuperscript{2}, F. Qureshi\textsuperscript{1}

\textsuperscript{1}Octave Bioscience, Inc., Menlo Park, CA, USA, \textsuperscript{2}Olink Proteomics, Uppsala, Sweden.
Background and Objective

- The availability of a validated blood-based assay to quantitatively assess disease activity and progression will significantly advance MS clinical care.

- There are two primary components to validate a test for commercial use in the clinic:
  - **Clinical Validation**: to establish that the test associates with endpoints relevant to assessing the disease. (Study in progress: see P014 ACTRIMS 2021)
  - **Analytical Validation**: to ensure that the test is characterized according to prespecified analytical performance criteria.

- One barrier to having a validated blood-based assay available to neurologists has been the lack of accurate, precise, and robust methods that ensure consistency in reported results over time.

- **Objective**: To analytically validate a blood-based multiplex proteomic immunoassay for MS disease assessments.
Methodology

- A custom immunoassay panel that measures the concentration of 21 proteins was developed using Proximity Extension Assay (PEA) methodology on the Olink™ platform.
- Selection of the 21 proteins was based on associations with clinical and radiographic MS endpoints in several independent discovery studies. Biological pathway modeling and protein network analysis were performed to ensure comprehensive representation of MS neurophysiology (see Chitnis et al., P0063-MS Virtual 2020).
- Two lots of the panel were manufactured for which the critical reagents (paired antibodies used for analyte binding, enzymes used for PCR steps, and protein stocks used for calibration) were varied to the extent possible.
- The two kit lots were subjected to a comprehensive analytical validation protocol to characterize and establish the following parameters in patient serum:
  - Accuracy, precision (inter- and intra-assay), robustness (instrumentation and analysis location), sensitivity (limits of quantification), MS population reference ranges, interference (endogenous interfering substances, heterophilic antibodies, common drugs and DMTs), diurnal variability, and stability of serum samples.

A multiplexed standard curve was prepared along with 3 levels of Calibrators (High Mid, Low) to cover the range of sample response in the MS population. The 4PL standard curve is re-established for each lot of reagent kits during manufacturing. The calibrators and a blank are run on each assay plate. The signal obtained from the assay (NPX) is converted to absolute concentration (pg/mL) using the calibrators referenced back to the "gold" standard curve. These concentrations are then used as inputs into algorithms corresponding to Disease Activity and Biological Pathway Scores.
Accuracy

- **Accuracy** is defined as the closeness of a result to the true value.
- 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 enables the accuracy assessment to be performed using endogenous protein (vs. a recombinant protein source).
  - Expected concentrations (i.e. the true values) were calculated by applying the targeted ratios of the unmixed samples.
- 4 individual samples were blended at various ratios:
  - Ratios of sample blends for mixtures with 2 samples were 25%:75%, 50%:50% & 75%:25%
  - Ratios of sample blends for mixtures with 4 samples were 25%:25%:25%:25% & 40%:10%:40%:10%
  - Samples were selected from an internal MS Cohort (n=64) to target both High and Low concentrations (relative to the MS population)
  - n = 20 mixed samples per each biomarker
  - Acceptability Criteria: Median % recovery for each biomarker must be between 80% - 120%. All analytes passed the specification.
### Inter and Intra-Assay Precision

**Precision** is defined as the extent to which repeated measurements agree with one another.

Intra-Assay Precision (within a single plate) and Inter-Assay Precision (across multiple plates) was characterized for each analyte. The percent coefficient of variation (%CV) was determined using serum pools enabling the assessment to be performed using endogenous protein.

- Equipment, reagents and location (i.e. R&D vs. Clinical Lab) were varied throughout the experiments to demonstrate the robustness of the method.
- The serum pools were manufactured to represent different populations including: two separate MS pools (PC MS 1 & 2 with shorter vs. longer disease duration), one rheumatoid arthritis pool (PC RA; an inflammatory disease control), and one healthy control pool (PC NM).
- These serum pools (n=4) were included on all R&D runs to date during the assay discovery and development process. They were sourced in large volumes, aliquoted, and stored at -65°C. They will also serve as the Process Controls (PC) used to assess acceptability of future analytical runs (run in triplicate on every plate). The standard deviation of repeated measurements is applied to the expected concentrations to create control tables for this purpose.

**Acceptability Criteria** was established as ≤15% for Intra-Assay %CV and ≤20% for Inter-Assay %CV.

- Intra-Assay Precision Experiment: 12 replicates per serum pool analyzed on a single plate
- Inter-Assay Precision Assessment: Up to 51 values per serum pool analyzed across 51 plates (spread across two manufactured kit lots)

**COL4A1** was found to have unacceptable inter and intra-assay precision across several serum pools. All other analytes passed the established criteria.

<table>
<thead>
<tr>
<th>Analyte</th>
<th>Intra &amp; Intra-Assay %CV with Expected Concentrations</th>
<th>Analyte</th>
<th>Inter &amp; Intra-Assay %CV with Expected Concentrations</th>
<th>Analyte</th>
<th>Inter &amp; Intra-Assay %CV with Expected Concentrations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC ID -&gt;</td>
<td>PC MS 1</td>
<td>PC MS 2</td>
<td>PC RA</td>
<td>PC NM</td>
</tr>
<tr>
<td>CCL20</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>6.9</td>
</tr>
<tr>
<td>CD6</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>5</td>
<td>6.8</td>
</tr>
<tr>
<td>CDCP1</td>
<td>9</td>
<td>8</td>
<td>9</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>CNTN2</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>8</td>
<td>6.8</td>
</tr>
<tr>
<td>COL4A1</td>
<td>6</td>
<td>7</td>
<td>6</td>
<td>8</td>
<td>6.8</td>
</tr>
<tr>
<td>CXCL13</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>6.8</td>
</tr>
<tr>
<td>CXCL9</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>8.9</td>
</tr>
<tr>
<td>FLRT2</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8.9</td>
</tr>
<tr>
<td>GFAP</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>9</td>
<td>9.2</td>
</tr>
<tr>
<td>GH</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>7</td>
<td>9.2</td>
</tr>
<tr>
<td>IL-12B</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8.9</td>
</tr>
<tr>
<td>MOG</td>
<td>6</td>
<td>9</td>
<td>6</td>
<td>8</td>
<td>8.9</td>
</tr>
<tr>
<td>NEFL</td>
<td>6</td>
<td>8</td>
<td>7</td>
<td>8</td>
<td>8.9</td>
</tr>
<tr>
<td>OPG</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8.9</td>
</tr>
<tr>
<td>OPN</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>9.2</td>
</tr>
<tr>
<td>PRTG</td>
<td>6</td>
<td>8</td>
<td>5</td>
<td>7</td>
<td>8.9</td>
</tr>
<tr>
<td>SERPINA9</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8.9</td>
</tr>
<tr>
<td>VCAN</td>
<td>6</td>
<td>8</td>
<td>9</td>
<td>7</td>
<td>8.9</td>
</tr>
</tbody>
</table>

- **Precision is defined as the extent to which repeated measurements agree with one another.**
- **Intra-Assay Precision** (within a single plate) and **Inter-Assay Precision** (across multiple plates) was characterized for each analyte. The percent coefficient of variation (%CV) was determined using serum pools enabling the assessment to be performed using endogenous protein.
  - Equipment, reagents and location (i.e. R&D vs. Clinical Lab) were varied throughout the experiments to demonstrate the robustness of the method.
  - The serum pools were manufactured to represent different populations including: two separate MS pools (PC MS 1 & 2 with shorter vs. longer disease duration), one rheumatoid arthritis pool (PC RA; an inflammatory disease control), and one healthy control pool (PC NM).
  - These serum pools (n=4) were included on all R&D runs to date during the assay discovery and development process. They were sourced in large volumes, aliquoted, and stored at -65°C. They will also serve as the Process Controls (PC) used to assess acceptability of future analytical runs (run in triplicate on every plate). The standard deviation of repeated measurements is applied to the expected concentrations to create control tables for this purpose.

- **Acceptability Criteria** was established as ≤15% for Intra-Assay %CV and ≤20% for Inter-Assay %CV
  - Intra-Assay Precision Experiment: 12 replicates per serum pool analyzed on a single plate
  - Inter-Assay Precision Assessment: Up to 51 values per serum pool analyzed across 51 plates (spread across two manufactured kit lots)

- **COL4A1** was found to have unacceptable inter and intra-assay precision across several serum pools. All other analytes passed the established criteria.
Sensitivity and MS Reference Ranges

- **Sensitivity** is defined as the assay's ability to accurately and precisely detect low concentrations of a given substance in a biological specimen.
- To establish the upper and lower limits of quantitation (LOQ), an LOQ panel was manufactured. For each analyte, 4 levels were targeted near the anticipated upper limit (ULOQ 1-4) and 4 levels near the anticipated lower limit (LLOQ 5-8). The targeted concentrations were based on the shape of the standard curve and location of asymptotes.
- This LOQ panel was run in triplicate over 2 lots (min. of 5 runs per lot) and fit to the standard curve. Accuracy (80%-120% recovery relative to expected concentration) and Precision (Inter-Assay CV ≤ 20%) assessments were used as the acceptability criteria to establish each analyte's LLOQ and ULOQ.
  - Serum samples that recover above the ULOQ or below the LLOQ will be imputed to the established LOQ level for reporting purposes.
- 928 MS serum samples were analyzed during the assay development process and used to establish MS Reference Ranges for each analyte. The linear interpolation method per Clinical Laboratory Standards Institute (EP28A3CE) was used to establish the 95% interval (2.5th and 97.5th percentiles).
  - For reporting purposes, the percentile relative to these reference ranges will be presented alongside the protein concentration.
  - To evaluate the impact of these established LOQs, the 928 samples were evaluated to determine the count and percentage of instances where the determined concentration was outside of the LOQ ranges. The assay is highly sensitive, with the maximum percentage of samples requiring imputation at any LOQ being 2.0% (CDCP1 at the LLOQ).
Assay Interference

Assay Interference: MS Therapeutics, Common Drugs, Heterophilic Antibodies and Endogenous Interferents

- **Assay interference** is defined as the effect of a substance present in the sample that alters the correct value of the result.

- To evaluate analytical interference of therapeutics in the assay 4 serum pools were spiked with MS DMTs and common drugs:
  - MS Therapeutics/Disease-Modifying Therapies (DMTs): Tested concentrations were targeted at 2X the reported $C_{\text{max}}$ (pharmacokinetic results from literature), or the highest possible concentration allowable for spiking with the procured interferent stock.
  - A universal monoclonal antibody standard was used as a surrogate for several DMTs and tested at two concentrations.
  - Components and concentrations for common drugs based on expert panel recommendations and Clinical Laboratory Standards Institute (CLSI EP-7A).

- For Rheumatoid Factor (RF) interference, RF concentrate (Lee Biosolutions) was used to spike 6 serum samples at high and low concentrations.

- For Human Anti-Mouse Antibodies (HAMA) interference, 5 HAMA positive serum samples were mixed at different ratios with MS samples from an internal cohort.

- For endogenous interference 4 serum pools were spiked with Bilirubin, Hemolysate, and Lipids at typical concentrations (Assurance Interference Test Kit).

- Median % Recovery was calculated relative to a corresponding spike control (represents the same alteration of the serum sample without the addition of the interferent) or for HAMA relative to the expected concentrations as determined from the unmixed samples.

- 80% to 120% median recovery across all spikes or sample mixtures was established as the acceptability threshold for the interference assessment.

  - For COL4A1 under recovery was observed for several drugs. Results are likely an artifact of established assay imprecision.
  - For CCL20, Cefoxitin spiked at 660 mg/dL resulted in median percent recovery of 77%. This finding will be further characterized.
  - For all other biomarker/interferent combinations, the median percent recovery was observed to be within 80 - 120%.
A Diurnal Variability study was performed to characterize fluctuations that can occur in biomarker levels between days over a relatively short duration. This initial study consisted of 8 subjects that had a serum sample collected at 6 time points: Day 1, Day 2, Day 3, Day 4, Day 5, and Day 12. For each timepoint & subject, the % difference of the observed protein concentration was calculated relative to the average concentration determined from all 6 timepoints. Additionally, the %CV was calculated for all 6 time points per patient.

Mean & median % differences for each biomarker/subject were observed to be within ± 20%. Mean and median % CV was found to be <20% for 18 of the 21 biomarkers. A follow up study restricted to the MS population including multiple draws within a single day and additional timepoints beyond 12 days is in the planning stages.

- A higher level of diurnal variability observed for COL4A1 is likely the result of the established assay imprecision.
- Mean and Median % CV for CCL20 (chemotactic cytokine) was observed to be >20%. Result will be characterized further in future studies with additional focus on relevance to MS disease endpoints.
- Growth Hormone (GH) results were found to be more variable than the other 20 biomarkers. GH is a biomarker that has been well established in the literature to have a high degree of ultradian and diurnal variability.
Sample Stability

Stability studies for serum samples have been performed to characterize storage and processing conditions anticipated in a clinical setting.

Stability was assessed at 4 temperatures: -65°C or below, -10°C or below (-20°C), 2-8°C (4°C), and room temperature (RT) using 4 MS serum samples.

- Results from -20°C, 4°C, and RT were compared to the control storage condition (-65°C or below) at the following timepoints:
  - 4 hours (for 4°C and RT only), Day 1, Day 3, Day 7, Day 14, and Day 28
  - All biomarkers were stable for up to 1 day at RT and 4°C and up to 28 days at -20°C
    - For room temperature: CXCL13, IL-12B, and TNFSF13B decreased beyond -20% at 3 days
    - For 4°C: OPN decreased beyond -20% at 3 days

- Additionally 5 Freeze-Thaw (FT 1-5) cycles (performed at the -65°C or below) were evaluated using 4 MS serum samples.
  - 3 Freeze-Thaw cycles were found to be acceptable using ± 20% difference (average) vs. the control condition (fresh sample) as the threshold.
    - GFAP concentrations decreased beyond -20% for Freeze Thaw cycles 4 and 5

Specifications for sample transport, processing, and storage will reflect results from this study and an ongoing expanded study (additional timepoints and samples).
Summary and Conclusions

- Performance has been assessed at the individual biomarker level. The custom assay panel has met acceptability criteria satisfying a fit-for-purpose analytical validation for 20 of the 21 proteins.
  - At present, COL4A1 has not met acceptability criteria due to assay imprecision.
- Additional validation experiments have been performed but not presented herein due to space constraints including: Cross-Reactivity (intra-panel and homologous proteins), Incurred Sample Reanalysis (ISR), and Plate Uniformity.
- Upon completion of the clinical validation study, the final Disease Activity and Disease Pathway Algorithms that utilize ensembles of proteins for the reported output will also be assessed for analytical validation parameters.
- A validated multivariate proteomic blood-based assay for objective MS disease assessments can serve as a quantitative, minimally invasive and cost-effective tool to enhance the standard of care for MS patients and their physicians.
  - The results of this analytical validation study will complement the ongoing clinical validation study and provide assurance that the assay has been thoroughly characterized and results are accurate, precise and robust.
- For more information, please contact:
  - Ferhan Qureshi: fqureshi@octavebio.com
  - Wayne Hu: whu@octavebio.com

Visit Additional ACTRIMS 2021 Octave Bioscience Posters: P014  P020  P025  P106