Abstract # **2409**

Analytical Validation of MSD S-PLEX® Neurology Panel 1

Catherine Demos[1], Nikhil Padmanabhan[1], Rachel Cohen[1], Jermaine Brown[1], Taron Gorham[1], Jack McBride[1], Sia Sesay[1], Sol Rivera Velez[2], Allan Barnes[3], Jun Yang[3], David Graham[2], Allen D. Everett[3], Christopher Campbell[1], Anu Mathew[1], Martin Stengelin[1], George Sigal[1], and Jacob Wohlstadter[1]

[1]Meso Scale Discovery, Rockville, Maryland, USA; [2]All Children's Hospital, Tampa, Florida, USA; [3]Johns Hopkins University, Baltimore, Maryland, USA

1 Abstract

Glial fibrillary acidic protein (GFAP), neurofilament light (Nf-L), and Tau are promising biomarkers for assessing neurological disorders and injuries such as hypoxic-ischemic encephalopathy (HIE) and Alzheimer's disease (AD). Detecting the low levels of these biomarkers in serum and plasma requires highly sensitive assays. Here, we report the characterization of a commercial ultrasensitive multiplexed immunoassay for GFAP, Nf-L and Tau, with the goal of analytically validating the panel for assessment of biomarker levels associated with HIE in infants.

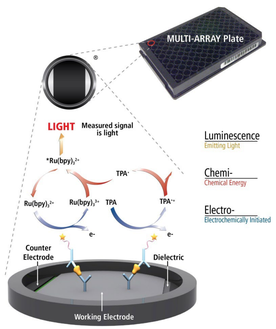
The MSD S-PLEX Neurology Panel 1 kit uses ultrasensitive S-PLEX assay technology to simultaneously measure levels of GFAP, Nf-L, and Tau (total) in a 96-well plate format. Analytical validation was performed in a series of studies based in part on Clinical and Laboratory Standards Institute (CLSI) guidelines to evaluate precision, dilution linearity, spike recovery, interference screening, and stability. Precision was assessed through a multi-site reproducibility study performed at MSD, Johns Hopkins University, and Johns Hopkins All Children's Hospital. Reagent stability was assessed through a two-year real-time stability study.

The quantifiable range of the assay was 2.4 – 4,200 pg/mL for GFAP, 7.6 – 10,000 pg/mL for Nf-L, and 0.34 – 740 pg/mL for Tau. A set of 15 common interfering substances was screened, and none showed interference exceeding 18%. Dilution of spiked and HIE samples up to 256-fold recovered at 80 – 120% of the expected value for all analytes. Precision error calculated using 40 samples across 3 independent laboratories and 3 production lots of the panel (360 measurements) was 18.2% for GFAP, 20.8% for Nf-L, and 17.5% for Tau. These commercially-sourced samples included adult and umbilical cord serum and plasma, challenge samples, and diseased samples with AD and other neurological disorders. Stability was measured with 20 plate runs per time point per lot at 12, 18, and 24 months; calculated drift was negligible, providing excellent shelf life for longer-term studies.

The MSD S-PLEX Neurology Panel 1 kit provides an analytically validated tool for assessing human GFAP, Nf-L, and Tau. The kit is sufficiently sensitive to detect these analytes in normal serum and plasma, the reproducibility and stability can support large studies, and its dynamic range is capable of quantifying elevated levels found in HIE and other target neurological disorders.

2 Methods

MSD's electrochemiluminescence detection technology uses SULFO-TAG™ labels that emit light upon electrochemical stimulation initiated at the electrode surfaces of MULTI-ARRAY® and MULTI-SPOT® microplates. The MSD® S-PLEX Neurology Panel 1 is a three analyte ultrasensitive panel. The S-PLEX platform uses ECL technology, retaining its well-known advantages and superior analytical performance. The improved sensitivity of S-PLEX assays is due to the proprietary TURBO-TAG® and TURBO-BOOST® reagents. This assay was validated using a custom protocol designed to reduce time to results. Incubation times are reduced by 30% compared to the commercial kit protocol and are performed at +27°C.



Electrochemiluminescence Technology

- Minimal non-specific background and strong responses to analyte yield high signal-to-background ratios.
- The stimulation mechanism (electricity) is decoupled from the response (light signal), minimizing matrix interference.
- Only labels bound near the electrode surface are excited, enabling non-washed assays.
- Labels are stable, non-radioactive, and directly conjugated to biological molecules.
- Emission at ~620 nm eliminates problems with color quenching.
- Multiple rounds of label excitation and emission enhance light levels and improve sensitivity.
- Carbon electrode surface has 10X greater binding capacity than polystyrene wells.
- Surface coatings can be customized.

3 Results

All analytical validation studies were designed according to CLSI guidelines where applicable. A modified, shortened assay method was used to return results within ~6 hours. Reproducibility estimates were generated in two stages: A single-site, single-lot precision study, then a multi-site, multi-lot precision study were designed and executed according to the CLSI Guideline EP05-A3. Quantification of the analytes was determined by reference to an eight point calibration curve (Figure 1) run on each plate and fit to a four parameter logistic (4PL) model. The limits of quantification were determined by measuring four high level, eight low level, and four blank samples in quintuplicate on three plates each for two reagent lots. The total error (TE) was calculated via the root mean squared model for each sample on each reagent lot and the TE cutoff value was ≤ 25%. (Table 1).

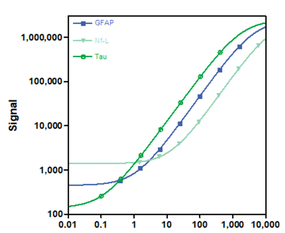


Figure 1: Representative calibration curves for each analyte in the MSD S-PLEX Neurology Panel 1 using the modified shortened protocol.

	Limits of Detection			
Conc. (pg/mL)	LOB	LOD	LLOQ	ULOQ
GFAP	1.72	2.05	2.40	4,200
Nf-L	3.23	4.99	7.65	10,000
tTau	0.255	0.324	0.34	740

Table 1: Limits of detection and quantification for the S-PLEX Neurology Panel 1 assay using the modified shortened protocol, dilution-adjusted.

4 Multi-Site Multi-Lot Precision Study

Three reagent lots were manufactured independently. Three sites were prepared with training, reagents, and identical sample source plates with 5 analytical controls (QC) and 35 samples in frozen, single-use aliquots. Each site performed 20 plate runs per lot over at least 20 calendar days. Standard curve and 5 QC sample concentrations were evaluated daily for acceptance criteria. Levy-Jennings plots and histograms were generated for each sample and analyte combination (Figure 2), then precision estimates were calculated for Site, Lot, Run, and Residual and Total %CV error. Estimates for all samples with mean concentrations within the limits of quantification were averaged for overall precision estimates for each assay (Table 2).

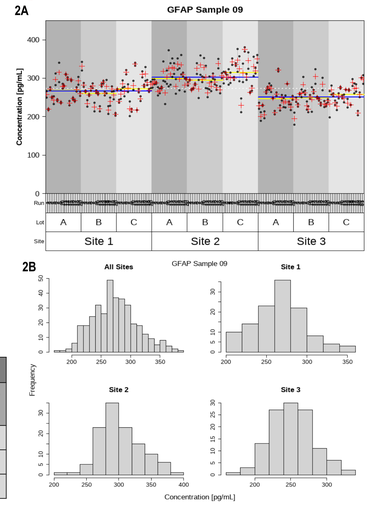


Figure 2: Levy-Jennings plot (A) and histograms (B) for GFAP measurements in Sample 09, umbilical cord serum spiked with calibrator, as a representative sample. Points are individual measurements, red crosses are mean values per Run, yellow lines indicate mean values per Lot within a given Site, blue lines indicate mean values per Site, and the dotted line indicates overall mean concentration. (B) Histograms show all measurements in the top left, then measurements filtered by Site.

Table 2: Overall precision estimates averaged across samples. Site and Lot are modeled as main effects, and Run is nested within Site-Lot combinations. Outliers were excluded by visual assessment and confirmed by MD68-statistic algorithm. Precision estimates were generated with R package VCA in R version 4.2.1.

	Table 2: Multi-Site Precision Estimates – %CV				
	Repeatability (Residual)	Between-Site	Between-Lot	Between-Run	Reproducibility (Total Error)
GFAP	8.6%	8.5%	4.4%	11.8%	**17.9%**
Nf-L	10.2%	7.9%	3.5%	13.1%	**19.4%**
Tau	7.5%	5.4%	3.9%	12.6%	**16.7%**

Challenge Sample Precision

Included among the 35 samples were challenge samples consisting of hemolyzed, human-anti-mouse antibody+ (HAMA+), or anti-nuclear antibody+ (ANA+) serum and plasma. Overall precision among challenge samples was comparable with overall precision for all samples (Table 3).

	Table 3: Challenge Sample Multi-Site Precision Estimates – %CV					
	Challenge	Repeatability	Between-Site	Between-Lot	Between-Run	Reproducibility
GFAP	HAMA+	8%	5%	4%	10%	**15%**
	Hemolysis	9%	5%	3%	12%	**16%**
	ANA+	9%	8%	4%	11%	**17%**
Nf-L	HAMA+	12%	4%	4%	15%	**21%**
	Hemolysis	10%	8%	3%	14%	**19%**
	ANA+	11%	8%	2%	15%	**21%**
Tau	HAMA+	7%	1%	5%	12%	**14%**
	Hemolysis	7%	3%	1%	12%	**15%**
	ANA+	9%	7%	4%	17%	**21%**

Table 3: Overall precision estimates for challenge samples. Site and Lot are modeled as main effects, and Run is nested within Site-Lot combinations. Outliers were excluded by visual assessment and confirmed by MD68-statistic algorithm. Each estimate is an average of one serum and one plasma challenge sample.

5 Single-Site Precision Study

We conducted a single site, single lot precision study to calculate precision estimates for Operator, Day, Run and total reproducibility within-laboratory. 40 plate runs were conducted by five operators over four days each (two plates/day staggered by at least two hours) asynchronously. Five analytical controls and ten samples were in duplicate on each plate. Variability was visualized with a Levy-Jennings plot (Figure 3).

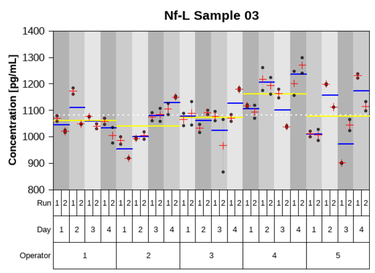


Figure 3: Levy-Jennings plot for Nf-L measurements in Sample 03 as a representative sample. Sample 03 is EDTA plasma spiked with calibrator. The Levy-Jennings plot groups by Operator, Day and Run, fully nested. Points are individual measurements, red crosses are mean values per Run, blue lines indicate mean values per Day per Operator, yellow lines indicate mean values per Operator, and the dotted line indicates overall mean concentration.

6 Real-Time Stability

To evaluate kit stability in real time, 20 plate runs conducted over at least 10 calendar days were tested at T=0, 12, and 18 months. The concentration of five analytical controls was plotted over time to determine any measurement drift (Figure 4). There was no change in the average measurements of the QC samples through 18 months post-manufacturing.

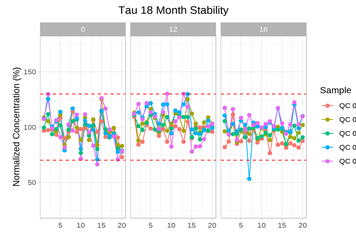


Figure 4: Real-time stability of the panel up to 18 months for five analytical control samples. Tau is shown as a representative of the panel. X-axis is Run number, facets are Timepoints in months, red dashed lines indicate +/- 30% of normalized concentration.

7 Dilution Linearity and Spike Recovery

Dilution linearity was verified using 64 individual commercially-sourced neonatal samples diluted 5- (recommended dilution for pediatric samples), 10- and 20-fold. Further dilutions were not tested due to the low sample volume available. Summary results for all samples are shown in Figure 5. GFAP, Nf-L and Tau are all observed to dilute linearly within 20%.

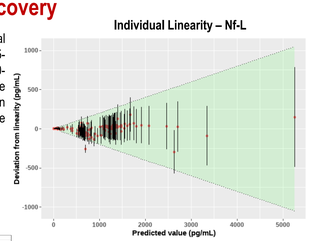
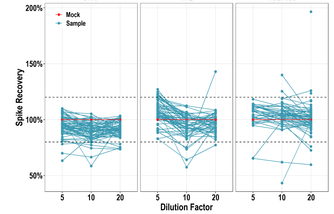


Figure 5: Summary of dilution linearity testing on 64 individual neonatal HIE samples for Nf-L. The shaded green area indicates the allowed limits of deviation (±20% from linear). Error bars indicate the 95% CI based on a typical intra-plate CV of 6%. All assays were observed to dilute linearly within 20%.

Figure 6: Spike recovery testing of 64 individual neonatal HIE samples (cord blood) showing recovery of GFAP, Nf-L, and Tau in most samples remains within ±20% (dashed lines) at 5-, 10-, and 20-fold dilution.



Spike recovery was determined first by spiking individual adult serum and plasma samples with three levels of calibrator, then diluting 2-fold (recommended dilution for adult samples). These GFAP, Nf-L and Tau assays are shown to recover the spiked calibrator within 80% – 20% at 2x dilution in serum and EDTA plasma, as shown in Figure 6. Next, individual neonatal HIE samples (n=64) were used to verify spike recovery. Samples were spiked with <5% overall volume of calibrator and diluted 5- (recommended dilution for neonatal samples), 10- and 20-fold as shown in Figure 6.

8 Interference Screening

Fifteen common laboratory and medical substances listed below were screened for interference with the MSD S-PLEX Neurology Panel 1. Samples spiked with the maximum clinically relevant concentration or a mock spike were compared, and none showed interference exceeding 18% in paired-difference testing. An example summary of interferent screening for the gentamicin effect on GFAP is shown in Table 4.

Table 4: Summary of interferent screening for gentamicin on GFAP. In all cases, the point estimate for the percent difference was below 18%, which leads us to conclude that gentamicin is not a significant cause of interference for GFAP at high or low concentrations. In most cases, the 95% confidence interval crossed 0, which indicates that the null hypothesis (i.e., gentamicin does not alter analyte measurements) could not be rejected.

Analyte	Sample Matrix	Analyte Level	Control Conc. (pg/mL)	Test Conc. (pg/mL)	Percent Difference (%)	Interference	95% CI
GFAP	EDTA Plasma	Spiked	1,070	1,030	-4.0	Not Detected	[-9.2,1.3]
	EDTA Plasma	Unspiked	16.0	16.8	4.4	Not Detected	[-1.8, 11]
	Serum	Spiked	737	735	-0.26	Not Detected	[-5.6, 5.1]
	Serum	Unspiked	20.9	20.6	-1.3	Not Detected	[-8.3, 5.6]

- Ampicillin
- Morphine
- Benzodiazapine
- Phenobarbital
- Morphine
- Erythropoietin
- Phenytoin
- Piperacillin
- Tazobactam
- Gentamicin
- Topiramate
- Levetiracetam
- Vancomycin
- Melatonin
- Bilirubin (conjugated)
- Bilirubin (unconjugated)

9 Conclusions

MSD has developed and analytically validated a highly sensitive, multiplexed immunoassay for quantifying GFAP, Nf-L, and Tau in plasma and serum. This assay is sufficiently sensitive to detect these analytes in normal serum and plasma and has a large dynamic range to accommodate elevated levels found in target neurological disorders. The panel has demonstrated excellent reproducibility across reagent lots, operators, and laboratories, providing a tool for decentralized and long term testing for GFAP, Nf-L and Tau in blood products.

DOWNLOAD POSTER

10 Funding

Research reported in this publication was supported by the National Institute of Neurological Disorders and Stroke of the National Institutes of Health under Award Number U01NS114144. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Meso Scale Discovery®
A division of Meso Scale Diagnostics, LLC.
www.mesoscale.com®

For Research Use Only. Not for use in diagnostic procedures.