**High-Likelihood ASD Infant Profiles: Enhancing Data Quality and Retention in Multi-site Early EEG Biomarker Studies**

**Authors:** Madison Booth, Abigail Dickinson, Adrian Lee, Bonnie Lau, Alana Campbell, Jed Elison, Juhi Pandey, John Zempel, Annette Estes, Stephen Dager, Heather Hazlett, Jason Wolff, Robert Schultz, Kelly Botteron, Natasha Marrus, Sara J. Webb, Neely Miller, Leigh MacIntyre, Alan Evans**,** Joe Piven, John R. Pruett, Jr. & Shafali Jeste for the IBIS Network

\*Current word count: 245 words/ 250 max

**Objective**: The multisite Infant Brain Imaging Study (IBIS) was established to identify early clinical and neurobiological markers of ASD. To enhance early detection and interventions, we examined pre-symptomatic markers of atypical development in infants who are at a higher likelihood of developing ASD (~20%; based on the presence of an older diagnosed sibling). To ensure clinical reliability, multisite biomarker studies require rigorous standardization of data collection, quality control, and signal processing. Here, we describe our standardization process and its impact on data quality.

**Methods:** At 6-months of age, all infants completed 1) the Vineland Adaptive Behavior Scale-III (VABS-III) and 2) EEG (task-free, visual, and auditory evoked potential paradigms). EEGs were standardized through rigorous behavioral management training, regimented recording procedures, and detailed post-processing data quality checks. EEG data was visually inspected for artifact free time periods and the proportion of data retained per infant for each paradigm was calculated.

**Results:** 6-month-old infants presented with variable degrees of cognitive and adaptive skills based on parent reports (Table 1). Despite early developmental differences, data retention rates were on average over 60% across all infants and all EEG paradigms (Table 2).

**Conclusion:** Previous studies report retaining infant EEG data from less than 50% of all participants and often do not report on quality of the data that are preserved (Cuevas et al., 2014). High quality EEG recordings are essential for understanding the relationship between progression of adaptive capabilities, clinical outcomes, and underlying neurodevelopmental biomarkers, all essential components in the establishment of early interventions.

**Table 1.** Mean Scores, Standard Deviations, and Minimum and Maximum Values for the Vineland-III across 42 6-month-old infants that also completed 6-month EEGs. VS = v-scale score

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Variable** | **Mean** | **SD** | **Min.** | **Max.** |
| **Communication STD Score** | 92.88 | 14.85 | 94 | 115 |
|  Receptive VS | 13.02 | 1.41 | 13 | 15 |
|  Expressive VS | 14.5 | 3.54 | 15 | 20 |
| **Daily Living Skills Standard Score** | 100.81 | 16.26 | 94 | 117 |
|  Personal VS | 15.12 | 2.83 | 14 | 18 |
| **Socialization Standard Score** | 104.76 | 4.95 | 100 | 107 |
|  Interpersonal Relationships VS | 15.4 | 0.71 | 14 | 15 |
|  Play and Leisure Time VS | 15.95 | 0.71 | 16 | 17 |
| **Motor Skills Standard Score** | 95.05 | 11.31 | 94 | 110 |
|  Gross VS | 13.86 | 2.13 | 11 | 17 |
|  Fine VS | 14.36 | 3.54 | 11 | 16 |
| **Adaptive Behavior Composite STD Score** | 98.62 | 5.22 | 103 | 106 |

**Table 2.** Mean percent of usable EEG data per paradigm across 83 6-month-old infants whose data has been analyzed since the incorporation of our extensive feedback protocols.

|  |  |
| --- | --- |
|  | **Average Percent of Usable Data Per Paradigm (%)** |
| **Data Collection Site** | **Resting** | **VEP** | **AEP** |
| Site 1 | 57.94±25.78 | 78.43±2.61 | 64.69±26.24 |
| Site 2 | 78.76±19.28 | 83.53±21.64 | 77.86±6.60 |
| Site 3 | 57.71±19.51 | 54.06±4.42 | 56.67±21.72 |
| Site 4 | 69.86±34.80 | 66.30±39.15 | 61.69±41.89 |
| Site 5 | 73.40±0.38 | 67.83±22.68 | 71.81±6.03 |