AUTOMATED PIPELINE FOR PREPROCESSING SCALP-RECORDED EEG DATA FOR PHASE-AMPLITUDE COUPLING ANALYSIS OF CHILDREN WITH AND WITHOUT INFANTILE SPASMS

Makoto Miyakoshi, Shaun A. Hussain, and Hiroki Nariai

Children's Discovery & Innovation Institute

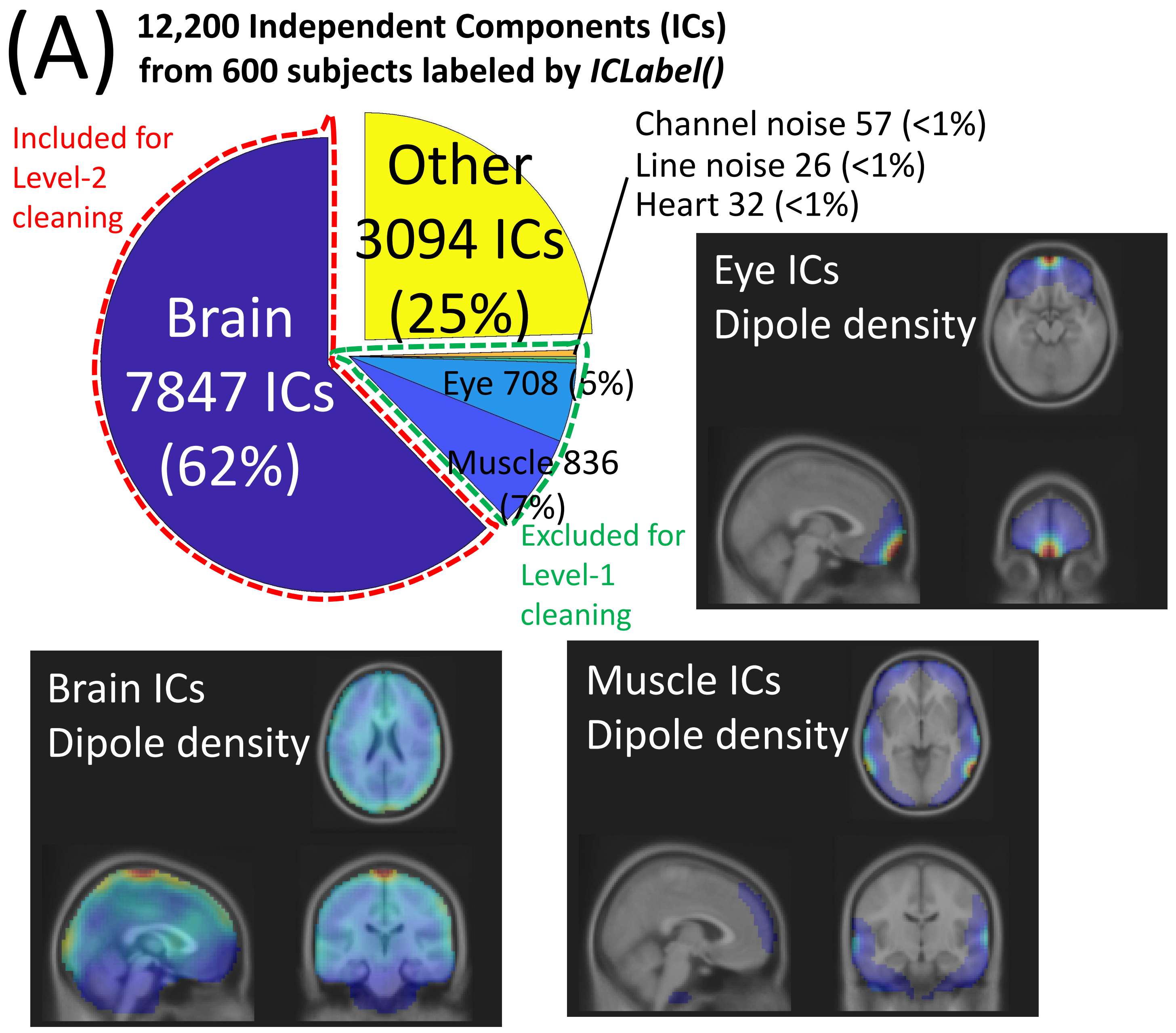
RATIONALE

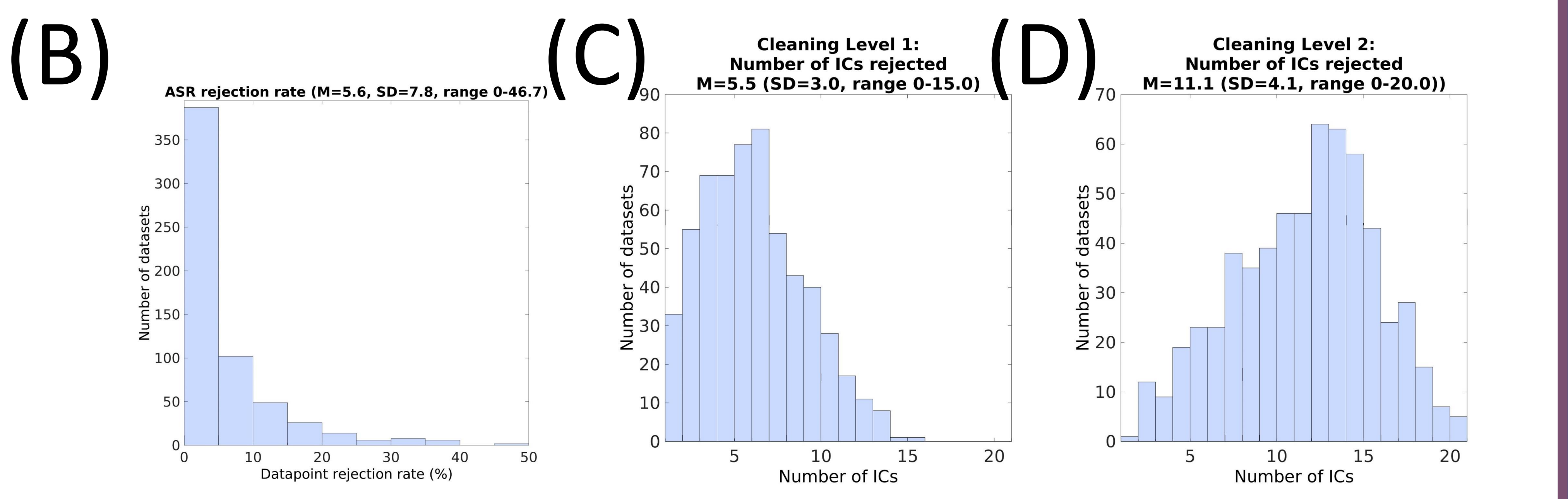
- Computational analysis of standard clinical EEG data is limited by abundant muscle, movement, and electrode artifacts, especially in the awake state. The mitigation of such artifacts is labor intensive and introduces methodologic risk, such as selection bias that may occur when a human electroencephalographer manually selects "artifact-free" samples for analysis.
- With goal of creating a stable preprocessing pipeline which can facilitate fully-automated processing of clinically-acquired EEG, we set out to design and validate a preprocessing algorithm using EEG data from children with epileptic spasms and normal controls, to obtain phase-amplitude coupling measures.

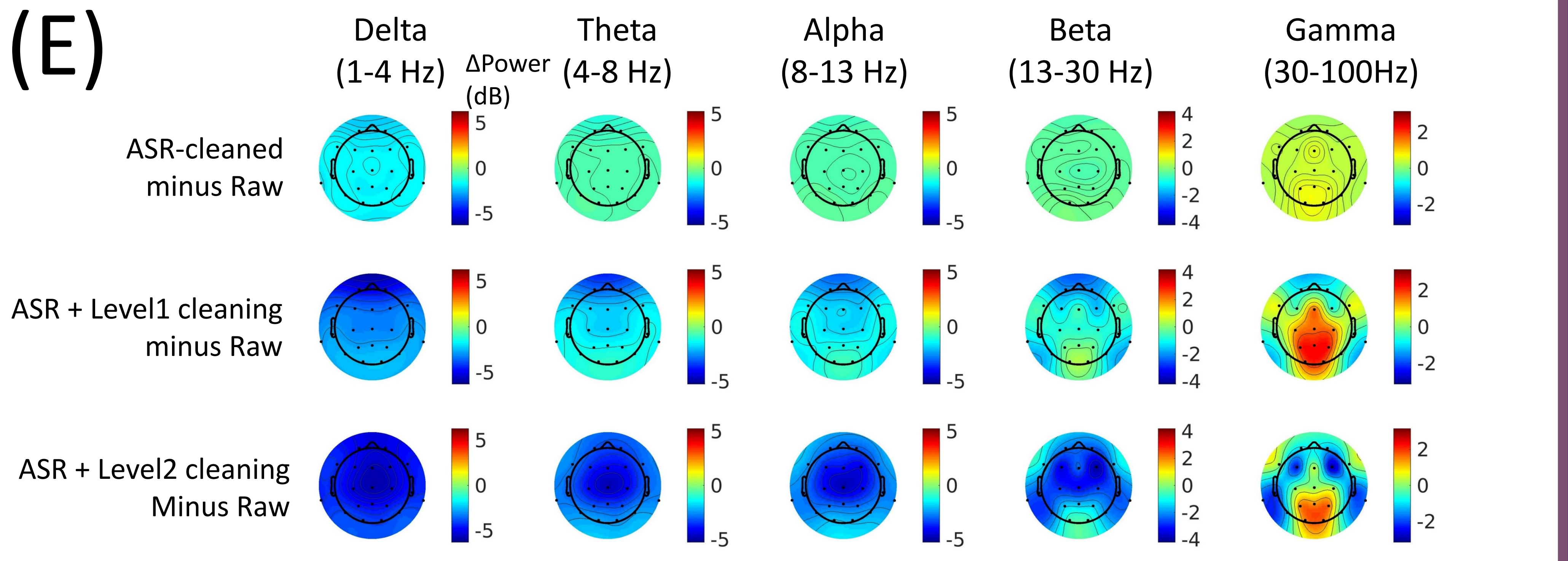
METHODS

- Total of 600 19-channel EEG samples recorded from both children with and without infantile spasms (cases and controls, respectively) were processed. These datasets were imported into EEGLAB14 [1] running under Matlab 2017b.
- The three PAC measures were obtained as follows: (1) After applying FIR high-pass filter at 0.5 Hz and *CleanLine()* for removing line noise, the first PAC measure (Canolty's modulation index, MI [2]) was calculated (phase 3-4 Hz, amplitude 35-70 Hz). (2) *clean_rawdata()* [3] was applied that included artifact subspace reconstruction (ASR), which is a solution for denoising for multivariate data using sliding-window PCA. Informax ICA [4] was performed to obtain independent components (ICs), which were evaluated with *ICLabel()* [5] to generate probabilistic labels.
- To evaluate the optimal level of data denoising, we prepared the three levels of cleaning: Raw, Level 1, and Level 2.
    - ➢ For the Level 1 cleaning, IC exclusion criteria were label probability > 0.8 for either Eye, Muscle, or Heart. After rejecting those ICs, the remaining ICs were backprojected to scalp electrodes for the second PAC calculation.
    - ➢ For the Level 2 cleaning, IC rejection was redone by selecting the class-label 'Brain', then the third PAC measure was calculated.
- As a validation procedure, we evaluated the number of ICs rejected and how well gamma-delta modulation indices (MI) could distinguish cases from controls using PACT [6].

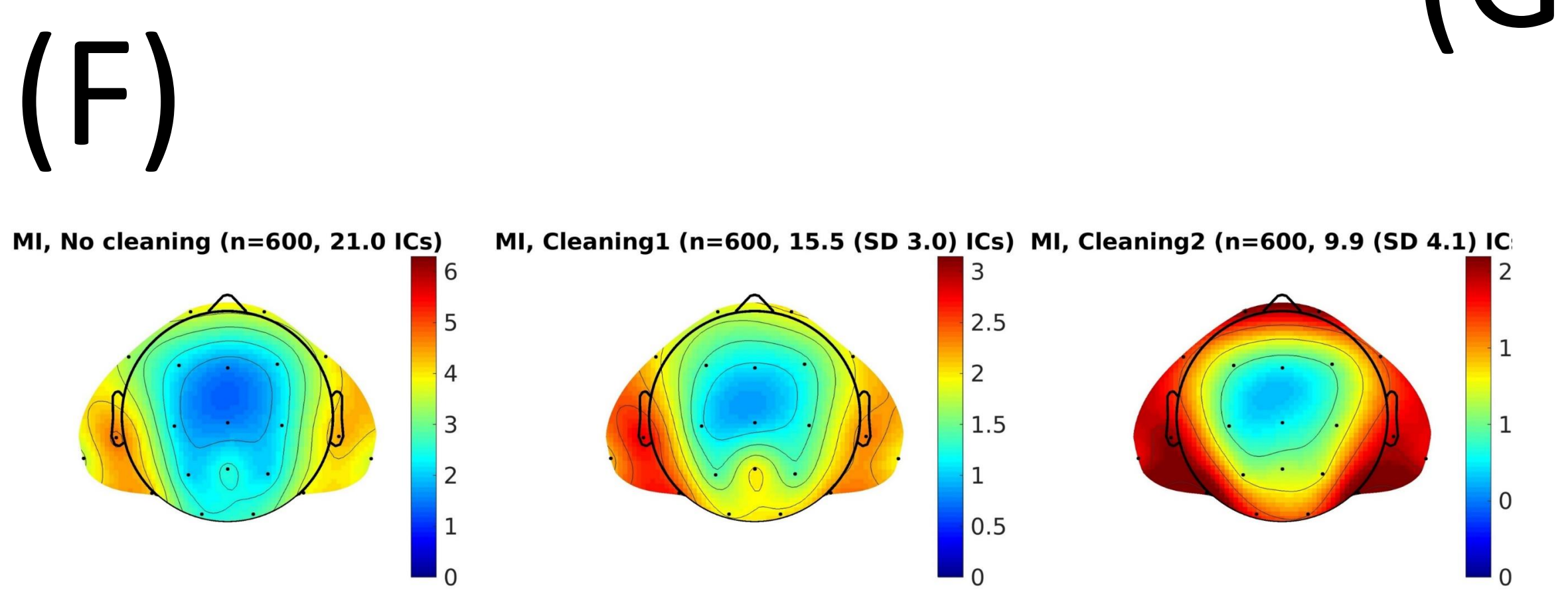
RESULTS

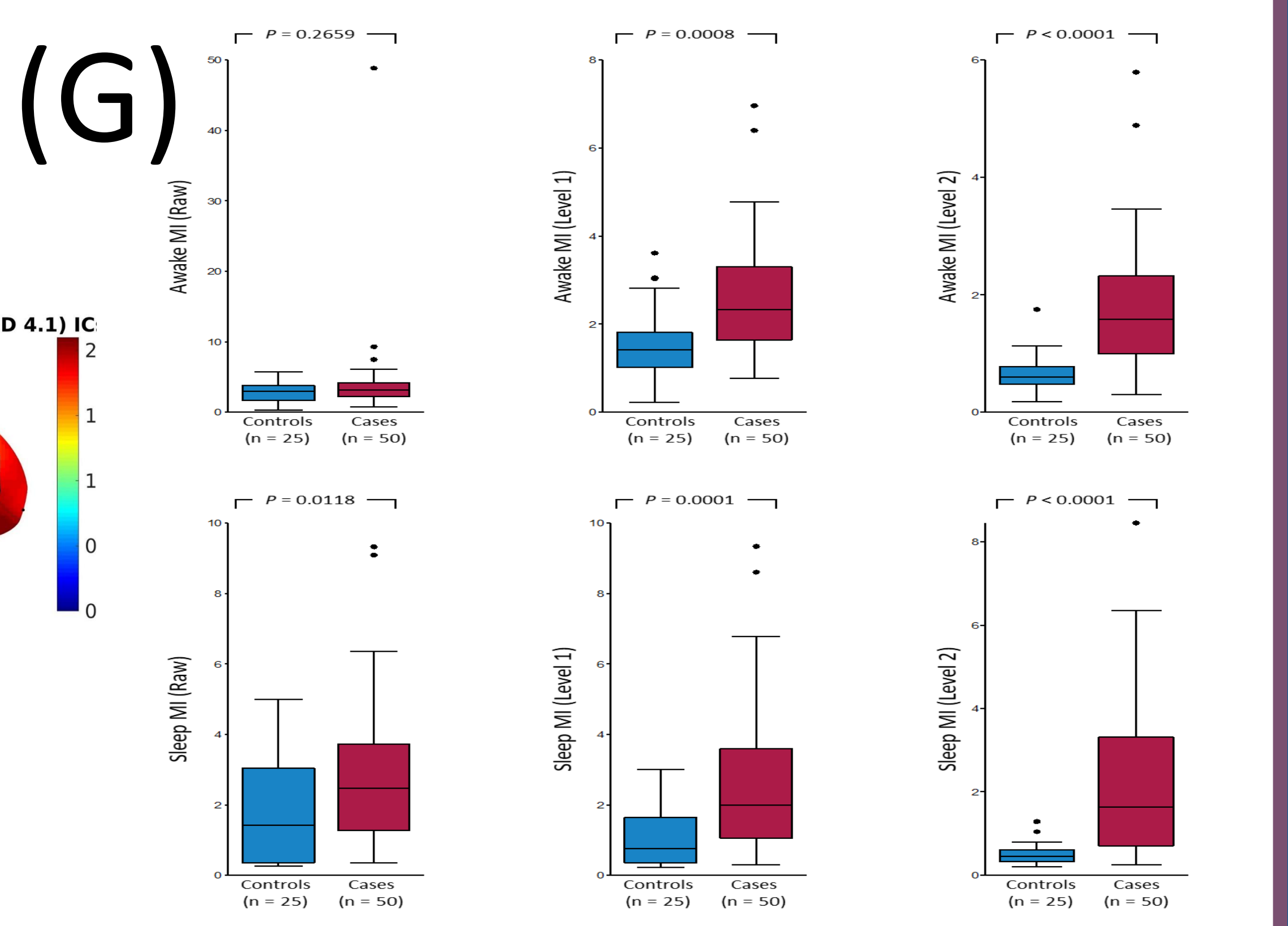






(A) 12,200 independent components (ICs) obtained from 600 subjects were labeled by ICLabel [5]. All the data were cleaned with artifact subspace reconstruction (ASR) [3]. For the level-1 cleaning, artifact ICs (highlighted in green) were rejected. For the level-2 cleaning, only brain ICs (highlighted in red) were included. For the brain-, muscle-, and eye-class ICs, corresponding dipole density plots are shown. **(B)** ASR removed average 5.6% of data points in addition to subspace reconstruction. **(C)** Level-1 cleaning rejected average 5.5/21 ICs. **(D)** Level-2 cleaning rejected average 11.1/21 ICs. **(E)** Effects of ASR and IC rejections compared with Raw data. For all scalp electrodes, EEG power differences were calculated for conventional EEG frequency bands. Note the opposite direction of the cleaning effects in the lower frequencies and higher frequencies. **(F)** Canolty's modulation index (MI) [2] calculated for different cleaning approaches. Note the color scales are different. **(G)** The effect of data cleaning. Both Level 1- and 2-cleaned data showed clearer differences between the control and clinical group.





DISCUSSION & CONCLUSIONS

We demonstrated that the fully automated preprocessing pipeline for the scalp-recorded EEG data can reveal clearer group difference between control group and the patients. Scalp electrode analysis revealed that both ASR and IC rejection globally lowers EEG power for the lower frequency ranges, but opposite effect was observed in the gamma range. Interestingly, the Level1 cleaning results showed larger gamma-power increase than that of Level2 cleaning, indicating inverse U-shape effect. The result was somewhat unexpected and warrants further investigation, as gamma-band power is directly related to calculation of modulation index.

Correspondence: Makoto Miyakoshi: mmiyakoshi@ucsd.edu
References: [1] Delorme and Makeig (2004) *J Neurosci Methods*, 134:9-21. [2] Canolty et al. (2006) *Science, 313*:1626-1628. [3] Chang et al. (2020) *IEEE Trans. Biomed. Eng.* 67:1114-1121. [4] Bell and Sejnowski (1995) *Neural Comput.* 7:1129-1159. [5] Pion-Tonachini et al. (2019) *NeuroImage*, 198:181-197. [6] Miyakoshi et al. (2013) *Conf Prof IEEE Eng Med Biol Soc*. 3282-3285.
Funding Support: This study was accomplished with support from the Swartz Foundation (Old Field, New York), the Susan S. Spencer Clinical Research Training Scholarship from the American Academy of Neurology, the Pediatric Victory Foundation, the Sudha Neelakantan & Venky Harinarayan Charitable Fund, the Elsie and Isaac Fogelman Endowment, the Mohammed F. Alibrahim Endowment, the Hughes Family Foundation, the John C. Hench Foundation, the UCLA Children's Discovery and Innovation Institute, and UCB Biopharma.