Using Quantitative EEG and Machine Learning to Predict Sevoflurane Concentration in Infants.

Rachit Kumar, BS[1] *; Georgia Georgostathi[2]; Olivia Nelson, MD[3]; Asif Padiyath, MD[3]; Nick Pratap, MD[3]; Allan Simpao, MD[3]; Fu-chiang Tsui, PhD[3]; Jack Wasey, BM BCh[3]; Shih Shan Lang, MD[4]; Ian Yuan, MD MEng[3 #] for the BRAIN collaborative investigators. * Presenting Author. # Corresponding Author.
Affiliations: 1: Medical Scientist Training Program, University of Pennsylvania; 2: School of Engineering and Applied Sciences, University of Pennsylvania; 3: Department of Anesthesiology and Critical Care Medicine, Children's Hospital of Philadelphia; 4: Division of Neurosurgery, Children's Hospital of Philadelphia.

INTRODUCTION:
Electroencephalography (EEG) monitors cortical electrical activity and can be used to assess depth of consciousness (DoC) during anesthesia. EEG waveforms require training to interpret and may be subject to interpreter bias. EEG can be processed with mathematical transformations that summarize EEG over a time epoch into an objective and replicable index—quantitative EEG (qEEG). In adults, proprietary qEEG (e.g., BIS index) has been used to prevent awareness under anesthesia and postoperative delirium.[1] However, these proprietary qEEG indices were not developed or validated in children, making them unreliable, especially in infants.[2] This knowledge gap is important since infants are particularly sensitive to the effects of excess anesthesia; 60% of infants ≤ 3mo experienced isoelectric EEG during anesthesia, an electrically inactive cortex associated with excess anesthesia and hypotension.[3] To assess the use of qEEG features to predict DoC in infants, we compared the accuracy of machine learning (ML) models applied to various qEEG features in predicting expired sevoflurane concentrations (eSevo) in infants ≤ 3mo.

METHODS:
EEG and eSevo in infants ≤ 3mo were extracted from a multicenter pediatric EEG study.[3] EEG were recorded on the Masimo Sedline EEG monitor. eSevo were recorded every minute and categorized into four levels: 0.1-1.0, 1.0-2.1, 2.1-2.9, and >2.9%. EEG from intubation to emergence were analyzed as one-minute epochs corresponding to one of four eSevo levels. EEG epochs were assessed for artifacts using NEURAL.[4] Epochs with ≥ 25% artifacts were excluded from analysis. Eligible epochs were processed with NEURAL to extract qEEG features: relative spectral power, connectivity coherence, and spectral entropy across frequency bands: 0.5-4, 4-7, 7-13, and 13-30Hz, in addition to burst percentage, spectral edge frequency 50 and 90%. The qEEG features for each minute epoch and corresponding eSevo levels were used to train eight ML models (Figure 1) using the scikit-learn package in Python. The dataset was randomly split 100 times into training and testing sets (80/20). After training on the training sets, the models' performances on the testing sets were averaged over the 100 iterations. Evaluation metrics include accuracy and F-measure, the harmonic mean of precision and recall.

RESULTS:
42 patients met inclusion criteria, representing 4619 epochs. eSevo levels 1-4 represented 6, 41, 32 and 21% of the data. Figure 1 presents the accuracy and F-measure of each ML model. Support vector machine (SVM), K-nearest neighbor (KNN), and Deep Multi-Layer Perceptron (DMLP) were the top performing models with median accuracies > 0.6 and F-measures > 0.5.

DISCUSSION:
In infants ≤ 3mo, SVM, KNN and DMLP offered higher accuracy in predicting eSevo levels based on qEEG features. Similar studies in adults, where BIS index was used as the gold standard, have yielded prediction accuracies of 89-95%[5,6]. In a study on infants ≤ 4mo, the prediction accuracy using SVM was 67-71%[7], similar to this study. The lower prediction accuracy in this study may be due to 1) infants having limited amplitude and frequency ranges compared to adults, resulting in less discriminatory power between eSevo levels; 2) limited EEG channels and data corresponding to the lowest eSevo level; and 3) lack of a gold standard (BIS) used in adult studies. Future direction will focus on using SVM, KNN, and DMLP on a reduced qEEG feature set, determined by computing *post hoc* feature importance.

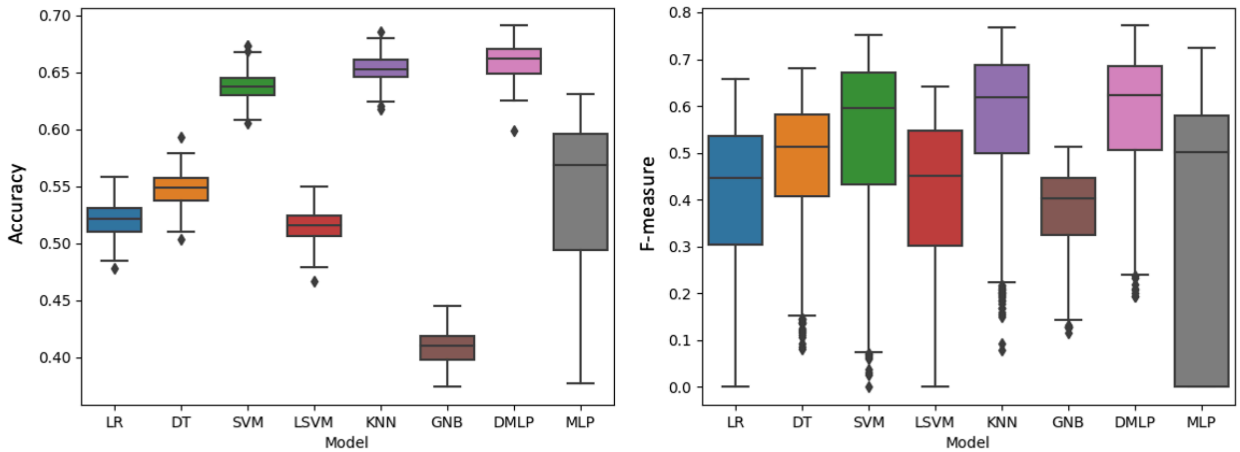

Figure 1: Boxplots of model accuracy (left) and F-measure with beta=1 (right) on holdout sequences over 100 different iterations.

LR = Logistic Regression; DT = Decision Tree; SVM = Support Vector Machine; LSVM = Linear Support Vector Machine; KNN = k-nearest neighbor; GNB = Gaussian Naïve Bayes; DMLP = Deep Multi-Layer Perceptron (with default Adam optimizer); MLP = Multi-Layer Perceptron (with stochastic gradient descent optimizer).

References:

1. Chan MT, Hedrick TL, Egan TD, et al. American society for enhanced recovery and perioperative quality initiative joint consensus statement on the role of neuromonitoring in perioperative outcomes: electroencephalography. Anesthesia & Analgesia 2020;130:1278-91.
2. Davidson A, Skowno J. Neuromonitoring in paediatric anaesthesia. Current Opinion in Anesthesiology 2019;32:370-6.
3. Yuan I, Xu T, Skowno J, et al. Isoelectric Electroencephalography in Infants and Toddlers during Anesthesia for Surgery: An International Observational Study. Anesthesiology 2022;137:187-200.
4. Toole JM, Boylan GB. NEURAL: quantitative features for newborn EEG using Matlab. arXiv preprint arXiv:170405694 2017.
5. Abel JH, Badgeley MA, Meschede-Krasa B, et al. Machine learning of EEG spectra classifies unconsciousness during GABAergic anesthesia. Plos one 2021;16:e0246165.
6. Gu Y, Liang Z, Hagihira S. Use of multiple EEG features and artificial neural network to monitor the depth of anesthesia. Sensors 2019;19:2499.
7. Chini M, Gretenkord S, Kostka JK, et al. Neural correlates of anesthesia in newborn mice and humans. Frontiers in neural circuits 2019;13:38.