Title: Diagnosing Pediatric Multiple Sclerosis from Visual Evoked Potentials with AI

Grace Hu, BSc, MD Candidate[1,2], Sahar Soltanieh, BASc, MASc[2], Chaojun Chen, BASc, ME[2], Rohit Sharma, RET, REPT[2], Farzad Khalvati, MASc, PhD[2,3] and Ann Yeh, MA, MD, FRCPC, Dip ABPN[2,4], (1)Temerty Faculty of Medicine, University of Toronto, Toronto, ON, Canada, (2)Neurosciences and Mental Health Program, SickKids Research Institute, Toronto, ON, Canada, (3)Department of Medical Imaging, University of Toronto, Toronto, ON, Canada, (4)Pediatric Neurology, University of Toronto, Hospital for Sick Children, Toronto, ON, Canada

Background: The optic nerve is affected in most individuals with multiple sclerosis (MS), with up to 1/4 of patients developing optic neuritis. Visual evoked potentials (VEP) are commonly used to evaluate optic nerve pathology in individuals with MS. These pathological changes are reflected in alterations to the morphology of VEP peaks, with patterns such as the "W" wave attributed to MS. Clinically, standardized thresholds based on p100 latencies are used to identify anomalous VEPs. Machine learning (ML) models may enhance the accuracy of current VEP-based morphological assessments in diagnosing MS.

Objectives: To develop and evaluate ML models that can distinguish between pediatric MS patients and non-inflammatory controls using VEP scans.

Methods: Consecutive children with MS (n=116, female=83) followed at the pediatric neuroinflammatory clinic at The Hospital for Sick Children and children with no neuroinflammation (n=23, female=12) received pattern-reversal VEP according to American Society of Clinical Neurophysiology Evoked Potential guidelines. P100 latencies were calculated per standard protocol from VEP scans.

For the ML assessment, scans were split into training (75%) and testing (25%) datasets, and SMOTE oversampling applied to balance the class distribution. Waveforms were processed through a ResNet1D encoder to extract 32 features, which were then used for the binary classification. Four ML models were used: Gradient Boosting Classifier, Random Forest Classifier, K-Nearest Neighbors Classifier, and C-Support Vector Classifier. Receiver Operating Characteristic Area Under the Curve (AUC) is used to evaluate the models' performance, with an unweighted average of class accuracies.

Results: 188 scans from MS patients (avg. age = 15.3, SD=2.5) and 28 scans from non-inflammatory controls (avg. age = 13.7, SD=3.8) were included in the study. Average p100 latency was 110.5 ms (SD=12.5) in MS patients and 106.2 ms (SD=9.3) in healthy controls. Using clinical thresholds of delayed p100 latency ≥115 ms in either eye or an inter-eye

latency difference ≥10 ms as the cutoff for anomalous VEPs, we were able to accurately identify 73.5% of non-inflammatory scans and 47.0% of MS scans.

Of the 4 models, Gradient Boosting Classifier performed best with an AUC of 0.62, achieving 42.9% accuracy for controls and 80.9% accuracy for MS. Other models showed varying AUC from 0.39 to 0.57. However, all models demonstrated limited performance in identifying healthy controls compared to MS patients.

Conclusions: ML models predict MS in children with greater accuracy than traditional methods based on p100 latency and inter-eye differences. Future studies will incorporate multimodal data to improve the accuracy of these ML models in diagnosing MS in youth.