

SEQUENTIAL MACHINE LEARNING MODELS FOR PREDICTING ONABOTULINUMTOXINA RESPONSE IN CHRONIC MIGRAINE: A COHORT STUDY

G. Vigani[1], M. Romozzi[2|3], A. Burgalassi[1], A. Lokhawandala[4], C. Vollono[2|3], P. Calabresi[2|3], F. Tascedda[5], A. Chiarugi[1|6], F. De Cesaris[1], L. Francesco Iannone[7]

[1]Headache Center and Clinical Pharmacology Unit, Careggi University Hospital, Florence, Italy; [2]Dipartimento Universitario di Neuroscienze, Università Cattolica del Sacro Cuore, Rome, Italy; [3]Neurologia, Dipartimento di neuroscienze, Organi di Senso e Torace, Fondazione Policlinico Universitario Agostino Gemelli IRCCS, Rome, Italy; [4]Drexel University, Philadelphia, USA; [5]Department of Life Sciences, University of Modena and Reggio Emilia, Italy; [6]Section of Clinical Pharmacology and Oncology, Department of Health Sciences, University of Florence, Italy; [7]Department of Biomedical, Metabolic and Neural Sciences, University of Modena and Reggio Emilia, Italy

Background. Chronic migraine (CM) is associated with high disability and heterogeneous response to onabotulinumtoxinA (BoNT-A). Baseline clinical predictors of response remain inconsistent, limiting individualized treatment decisions. Machine learning may improve prediction by integrating longitudinal treatment data. This study aimed to develop sequential machine-learning models to predict BoNT-A response in CM after successive treatment cycles under real-world conditions.

Methods. We performed a post-hoc analysis of a prospectively collected cohort of adults with CM treated with BoNT-A according to the PREEMPT protocol at a tertiary headache center. Response was defined as a ≥30% reduction in monthly headache days (MHDs) from baseline. Four sequential supervised classification models were trained to predict response after cycles 1-4. Model 1 used baseline demographic and clinical variables only; models 2-4 progressively incorporated post-treatment MHDs and acute medication use after preceding cycles. Multiple algorithms were compared using stratified 5-fold cross-validation, with balanced accuracy as the primary performance metric. Model performance was assessed using accuracy, sensitivity, specificity, positive and negative predictive values, macro-F1 score, and AUC-ROC. Feature importance was evaluated using SHapley Additive exPlanations.

Results. The baseline cohort included 385 patients with CM; mean age was 48.1±13.1 years, 82.9% were female, baseline MHDs were 25.4±5.9, and medication overuse was present in 326/383 patients (85.1%). The number of patients included decreased across models with 385 in model 1 and 161 in model 4. Baseline-only prediction showed limited performance, with balanced accuracy of 0.60±0.03 and AUC-ROC of 0.62±0.04. Incorporating early post-treatment data markedly improved discrimination. Model 2 achieved balanced accuracy of 0.82±0.04 and AUC-ROC of 0.88±0.01, while model 3 showed the highest balanced accuracy, 0.84±0.04, and AUC-ROC of 0.87±0.03. Across longitudinal models, on-treatment MHDs were the dominant predictors of response.

Conclusions. Sequential machine-learning models incorporating early post-treatment headache-frequency data outperformed baseline-only prediction of BoNT-A response in CM. These findings support dynamic, monitoring-informed prediction models and require prospective external validation before clinical implementation.