

P0205 Exploratory study on the application of digital natural language processing techniques for the classification of headache disorders

N. Vandenbussche^{1,2,*}, K. Plevoets^{3,4} and K. Paemeleire^{1,2}

1 Ghent University Hospital, Department of Neurology, Ghent, Belgium 2 Ghent University, Department of Basic and Applied Medical Sciences, Faculty of Medicine and Health Sciences, Ghent, Belgium 3 Ghent University, Faculty of Arts and Philosophy, Ghent, Belgium 4 Ghent University, Faculty of Sciences, Ghent, Belgium

Background and objective: To investigate natural language processing (NLP) to classify between migraine and cluster headache based on patient-written text. Methods: Outpatient migraine and cluster headache patients were recruited to provide a written, digital, Dutch narrative text about their headache disorder. We analysed texts through manual annotation for themes, lexical and sentiment analysis. Machine learning (ML) models aimed to classify between both disorders based on attack description. The study was approved by the ethical committee of Ghent University Hospital. Results: One hundred twenty-one patients, 81 migraine patients (79% female, mean age 43) and 40 cluster headache patients (20% female, mean age 49), participated. Themes with the highest coverages were medical history, attack description, treatment and burden of disease. Word keyness in texts was significant for “hoofdpijn” (headache) in migraine and “pijn” (pain) in cluster headache (both $p < 0.001$). A trend towards higher negative emotional tonality in attack descriptions by cluster headache patients was found. ML models showed best results for naive Bayes classifiers (average [avg] accuracy [acc] 0.90, avg F1-score 0.85), compared to support vector machines (avg acc 0.80, avg F1-score 0.69) and logistic regression (avg acc 0.81, avg F1-score 0.63). Conclusions: NLP and ML applications have a high potential to classify between migraine and cluster headache based on patient-written attack descriptions