Title: A Machine Learning Approach to Predict Change in Diagnostic Category in Pre-Dementia

Sascha Gill[1,4], Pauline Mouches[1,2], Sophie Hu[1,3,4], Muhammad Ahmad[1], Nils D. Forkert [1,2], Zahinoor Ismail[1,4,5]

[1] Hotchkiss Brain Institute, University of Calgary, Calgary, AB, Canada
[2] Department of Radiology, University of Calgary, AB, Canada
[3] Department of Community Health Sciences, University of Calgary, AB, Canada
[4] Mathison Centre for Mental Health Research & Education, University of Calgary, Calgary, AB Canada
[5] Department of Psychiatry, University of Calgary, Calgary, AB, Canada

Methodological Question: Are neuropsychiatric symptoms (NPS) categorized into Mild Behavioural Impairment (MBI) domains important features to predict change in diagnostic category in pre-dementia risk states?

Introduction:
Alzheimer's Disease (AD) is the most common form of dementia in elderly people worldwide, with over 560,000 individuals in Canada currently living with AD. Mild Cognitive Impairement (MCI), is a prodromal stage of AD associated with an estimated annual progression rate of 10-15%. Classification methods have been developed to predict diagnosis of MCI and AD based on different modalities of biomarkers, e.g., clinical, imaging, genetic, and cerebrospinal fluid markers etc. However, the presence of NPS has not been explored in these predictive models. NPS such as apathy, mood, anxiety, lack of impulse control, disinhibition, and psychosis in older adults and MCI are associated with higher risk for cognitive decline and dementia. MBI is a neurobehavioural syndrome that describes later-life onset of NPS as an at-risk state for cognitive decline. MBI symptoms are divided into five domains including drive/motivation, emotion dysregulation, impulse dyscontrol, social cognition, and abnormal perception. The aim of this study is to use machine learning to identify features essential for predicting change in diagnosis (i.e., NC-to-MCI-to-AD; or MCI-to-AD progression) using baseline data from the Alzheimer's Disease Neuroimaging Initiative (ADNI) study.

Methods: Baseline neuroimaging, neuropsychiatric, and clinical data from 108 NC and 275 MCI patients were extracted from the ADNI database to train the classifier to distinguish between individuals who deteriorated in their diagnosis and those who did not. Final clinical diagnosis at the latest follow-up was recorded. More than 200 features were considered as potential predictors of progression. The performance of several feature selection methods combined with different machine learning algorithms, including nearest neighbours, random forest, support vector machine, multilayer perceptron, and decision tree, was compared. The predictive utility was assessed using a 10-fold cross validation framework.

Results: The random forest algorithm, combined with a relief feature selector was adopted to evaluate the classification accuracy. The data was approximately balanced and included 23%

NC, 35% MCI and 42% AD at final diagnosis. The best performing model incorporated a combination of neuropsychiatric markers and MRI measures and predicted progression with 60.1% accuracy (0.6 sensitivity, 0.8specificity). A small subset of features, n=21 were selected to be important for the machine learning classification model. Predictors of progression included volume/cortical thickness of brain regions (left superior temporal, left hippocampus, left entorhinal) and MBI total, emotion dysregulation and impulse dyscontrol scores among others. This is a preliminary analysis, and future work will explore if the predictive accuracy can be improved with demographic and clinical characteristics such as cognitive scores.

Conclusions:
This study highlights the importance of neuropsychiatric symptoms in predicting change in diagnosis. Prognostic classification of NC and MCI at the individual patient level has the potential to improve clinical trial design, identify patients for early treatment, as well as guide clinical and patient decision-making. Utilization of the MBI-checklist can improve in early detection and treatment of individuals at high-risk pre-dementia states.