[BIOINFO 2023]

Prediction of Alzheimer's disease using AI model based on SNP chip data

Myeongji Cho[1], Hye-Ryeong Nam[1], Ju-Hee Lee[1], and Sang Cheol Kim[1,*]

[1]*Division of Healthcare and Artificial Intelligence, Department of Precision Medicine, National Institute of Health, Korea Disease Control and Prevention Agency, Cheong-Ju, Republic of Korea*
**Corresponding author: sckim.knih@korea.kr*

Alzheimer's disease (AD) is a complex neurodegenerative disorder involving multiple pathogenic factors. Mild cognitive impairment (MCI), a condition between normal and dementia, can be considered an early stage of AD, and approximately 10-15% of MCI patients convert to AD each year. Thus, early detection of cognitive impairment in the early stages of AD and timely intervention can effectively delay the onset of AD. In this study, we conducted a study to predict cognitive function changes through artificial intelligence-based predictive modeling using random forest (RF), k-nearest neighbor (KNN), artificial neural network (ANN), and support vector machine (SVM) algorithms. We used data from participants in the Biobank Innovations for chronic Cerebrovascular disease With ALZheimer's disease Study (BICWALZS) at Ajou University Hospital in South Korea. GWAS was performed using Korean Chip-based genotype data on 674 Koreans to obtain 767,012 SNPs associated with dementia, which were used to construct training matrices for prediction models. Quality control (QC) procedures included (1) call rate check, (2) hetero rate check, (3) multidimensional scaling, (4) singleton check, and (5) removal of SNPs with more than 5% missing values, minor allele frequency smaller than 1%, and Hardy-Weinberg equilibrium p value < 10^{-6}. Using genotype data of healthy controls and dementia patients, we trained and evaluated the prediction model, and applied the constructed model to 80 MCI participants with 2-years follow-up information to compare with the actual changes in cognitive function status. All models predicted cognitive decline in four with AD out of the five individuals who actually developed cognitive decline among 80 MCI participants. The remaining one, whose cognitive decline was not predicted, was confirmed to be a vascular dementia (VD) patient, suggesting that further research is needed to utilize VD-specific genetic information. In addition, the high false positive rate of the model may imply the possibility of cognitive changes in MCI subjects in the future, so longer follow-up data is needed.