



Application number/Title: 33022: Development of a computational method to predict disease risk

Applicant PI: Jinho Kim

Application Institution: Samsung Medical Center, South Korea, Seoul

Keywords provided by the Applicant PI to describe the research project:

Biomarker, Disease Risk Prediction, Machine Learning

Application Lay Summary:

The aim of the proposed research is developing a new computational method which enables to predict disease risk of an individual based on the genetic and environmental factors. Numerous GWA study demonstrated that genetic factor can predict the onset of some disease. Also, the incidence of disease changes depending on the lifestyle and the exposed environment. We will use advanced machine learning techniques to build models that predict disease risk determined by genetic and environmental factors and their interactions. Since environmental factors influence the expression of genotypes, it is necessary to consider environmental factors together for accurate modeling. The predictive models we are building can be used for predicting disease risk using the genetic and environmental factors collected in the BioBank to find key factors for the disease risk. We will develop computational algorithms that predict disease risk for healthy people. The UK Biobank resource is practically significant in its ability to support our research. As a first step, we will determine the genetic factors that affect the disease. As a second step, we will use the artificial intelligence (AI) learning methodology to learn the complex relationships that these factors have on the incidence of disease. We will test various kinds of AI methods to select the method with the most predictive accuracy. In this step we will consider environmental variables together as predictors. Finally, we will calibrate the model predictions to the population data so that absolute disease risk can be calculated. We expect it will take three years to complete these tasks. These models can provide the most appropriate preventive measures to the people with such diseases. And it will produce the greatest effect at the minimum cost. We also expect that this model will improve the quality of life of individuals by distinguishing the possibility of occurring diseases gradually so to select which disease has to be cured at first.