

# Conditional Quantile Random Forest with its application for predicting the risk (Post-Traumatic Stress Disorder) PTSD after experienced an acute coronary syndrome

Huichen Zhu

*The Hong Kong University of Science and Technology*  
*E-mail: hz2366@cumc.columbia.edu*

**Abstract:** Classification and regression trees (CART) are a classic statistical learning method that efficiently partitions the sample space into mutually exclusive subspaces with the distinctive means of an outcome of interest. It is a powerful tool for efficient subgroup analysis and allows for complex associations and interactions to achieve high prediction accuracy and stability. Hence, they are appealing tools for precision health applications that deal with large amounts of data from EMRs, genomics, and mobile data and aim to provide a transparent decision mechanism. Although there is a vast literature on decision trees and random forests, most algorithms identify subspaces with distinctive outcome means. The most vulnerable or high-risk groups for certain diseases are often patients with extremely high (or low) biomarker and phenotype values. However, means-based partitioning may not be effective for identifying patients with extreme phenotype values. We propose a new regression tree framework based on quantile regression that partitions the sample space and predicts the outcome of interest based on conditional quantiles of the outcome variable. We implemented and evaluated the performance of the conditional quantile trees/forests to predict the risk of developing PTSD after experiencing an acute coronary syndrome (ACS), using an observational cohort data from the REactions to Acute Care and Hospitalization (REACH) study at New York Presbyterian Hospital. The results show that the conditional quantile based trees/forest have better discrimination power to identify patients with severe PTSD symptoms, in comparison to the classical mean based CART. This is joint work with Ying Wei, Katharina SchulteBraucks and Ian Kronish.