

ABSTRACT

THESIS: Subgroup Identification for Differential Treatment Effect: Model-based

Recursive Partitioning Approach

STUDENT: Md Yasin Ali Parh

DEGREE: Master of Science

COLLEGE: Sciences and Humanities

DATE: July 2020

PAGES: 66

The goal of this study is twofold: i) identification of risk factors associated with three chronic diseases and with mortality under three scenarios, and (ii) identification of subgroups with differential treatment effects. Multivariate analysis is performed to identify the risk factors associated with chronic diseases, hypertension, diabetes, and dyslipidemia. Both multivariate parametric and semi-parametric regression models are applied to identify risk factors for time to death from all-cause mortality, from cardiovascular diseases (CVD) and from cancer. For subgroup identification, we applied a model-based recursive partitioning approach. This method fits a local model in each subgroup of the population rather than fitting one global model for the whole population. The method starts with a model for the overall effect of treatment and checks whether the overall impact of the treatment is equally applicable for all individuals under the study based on parameter instability of M fluctuation test over a set of partitioning variables. The procedure produces a segmented model with a differential effect of treatment corresponding to each subgroup. The subgroups are linked to predictive factors learned by the recursive partitioning approach. The methods are applied to the data from the Ball State Adult Fitness Program Longitudinal Lifestyle Study, where we considered the level of cardio-respiratory fitness (CRF) as a treatment variable. The overall results indicate that CRF is inversely associated with chronic diseases and mortality under three scenarios. The predictive factors that are selected for both chronic diseases and mortality scenarios in subgroup analysis are related to diseases and mortality scenarios. The subgroup-specific results of chronic diseases indicate that for each subgroup, the chance of chronic diseases increases with low CRF. The subgroup-specific results of all-cause mortality demonstrate that the risk of death for younger (age ≤ 49 years) is higher if their CRF is low. A similar scenario is observed for the subgroup-specific result of mortality due to CVD.