BREAST DISEASE RISK FACTORS

E-Mail: [moghbaie@stevens.edu](mailto:moghbaie@stevens.edu)

11/28/2016

Mehrnoosh Oghbaie

Stevens Institute of Technology

We are going to study a database contain information about the risk factors associated with benign breast disease. The dataset includes 200 observations and 14 variables. The data is comprised of a subset of large study in which the original data are from a hospital based case-control study designed to examine the epidemiology of fibrocystic breast disease. Data are provided on 50 women who were diagnosed as having benign breast disease and 150 age matched controls, with three controls per case. Matching was based on the age of the subject at the time of interview. Cases included women with a biopsy-confirmed diagnosis of fibrocystic breast disease identified through two hospitals in New Haven, Connecticut. Controls were selected from among patients admitted to the general surgery, orthopedic, or otolaryngologic services at the same two hospitals.

The independent variables are: (STR) stratum, (AGMT) age of the subject at interview, Degree, (CHK) regular medical checkups, (AGP1) age at first pregnancy, (AGMN) age at menarche, (NLV) # of stillbirths/ miscarriages, (LIV) number of live births, (WT) weight of the subject, (AGLP) age at last menstrual period, (MST) marital status.

The dependent variables are: (OBS) observation within a stratum, (FNDX) final diagnosis.

Independent variables form the latent variables, and final diagnosis would be the outcomes for factor analysis, logistic regression and linear regression. The predictors for logistic and linear regression will be determine after factor analysis.

# Project Proposal