

Abstract:

The aim of this study is to predict the risk that patients in primary care to develop end-stage renal disease (ESRD), using routinely collected electronic health records. Such predictions enable early detection of patients with high risks of ESRD and appropriate referral for specialized treatment. Data are obtained from the Salford Integrated Records, which is a database in the city of Salford, UK. The data include repeated serum creatinine measurements, time and event indicator for ESRD and a number of baseline explanatory variables. There are 35,001 patients who were followed between January 1, 2001 and December 31, 2014. The number of repeats per patients ranged between 1 and 345, with a mean of 15. We consider joint modelling of longitudinal and survival data. The

longitudinal sub-model is a linear mixed effects model, where we de-compose creatinine measurements (in log-scale) into fixed effects, a random intercept, an integrated random-walk and measurement error.

The survival sub-model is a Cox model with time-varying frailty. Longitudinal and survival data are linked by including an error-corrected version of the creatinine measurements and its instantaneous rate of change in the survival sub-model. Examples of dynamic predictions are presented.