Available online at www.sciencedirect.com

ScienceDirect

Procedia Environmental Sciences 27 (2015) 119 – 122

Procedia Environmental Sciences

Spatial Statistics 2015: Emerging Patterns

Predictive performance of geoadditive survival models to study geographical patterns in coronary heart disease

Ipek Guler[a], Jenifer Espasandín-Domínguez[a], Mar Rodríguez-Girondo[b], Francisco Gude[c], Carmen Cadarso-Suárez[a]*

a Unit of Biostatistics, School of Medicine, University of Santiago de Compostela, Spain
b Department of Medical Statistics and Bioinformatics, Leiden University Medical Center, Netherlands
c Clinical Epidemiology Unit, Hospital Clínico Universitario, Santiago de Compostela, Spain

Abstract

Acute coronary syndrome (ACS) represents the most common cause of death in the western world. Numerous prediction models exist for the different types of ACS. Most of these models have been developed from large populations by means of the classical (parametric) Cox proportional hazard model, in which the geographic area has not been taken into account as a health determinant. However, this statistical Cox model may not be enough to capture some flexible effects of covariates on survival, and does not allow to include spatial effects. In this study, we used flexible extensions of the Cox model, such as Structured Geoadditive Survival Models, to evaluate geographical inequalities in survival of patients admitted to a tertiary hospital, with a diagnosis of ACS. The predictive performance of the survival models were assessed through time-dependent Receiver Operating Characteristic (ROC) curves computed by the incident sensitivity and dynamic specificity for each time point.

Peer-review under responsibility of Spatial Statistics 2015: Emerging Patterns committee

Keywords: Structured Additive Regression; Penalized splines; sociodemographic effects; Cox model.

* Corresponding author.
E-mail address: ipek.guler@usc.es

Peer-review under responsibility of Spatial Statistics 2015: Emerging Patterns committee
doi:10.1016/j.proenv.2015.07.120

1. Introduction

Acute coronary syndrome (ACS, O'Connor et al. [1]) is a term for a condition due to reduced blood flow to the coronary arteries where the heart muscle is unable to function properly or dies.

There exist several methods to study the different types of ACS. However, in most of them, the geographic area has not been taken into account as a health determinant. Diverse scientific studies have shown that significant health inequalities exist and in many cases cause a higher mortality or sickness rate than most known health risk factors. Also in the medical literature, hypotheses to explain neighbourhood effects on ischemic heart disease often refer to the physical environment and socioeconomic factors. Therefore, the classical parametric Cox proportional hazard models may not be enough to study such kind of effects. In this paper, we use a flexible extension of the Cox models, Structured Geoadditive Survival models (Kneib and Fahrmeir [2]), in which flexible and geographical covariate effects to study the survival process of the patients with a diagnosis of ACS can be included.

The outline of the paper is as follows: the ACS database is explained in Section 2, the Structured Geoadditive Survival models are presented in Section 3, the results are discussed in Section 4 with the predictive performance of survival models using time-dependent ROC (Receiver Operating Characteristic) curves computed by the incident sensitivity and dynamic specificity for each time point.

2. Database: Acute Coronary Syndrome (ACS)

The objective of the study is to detect geographical inequalities in survival of patients with ACS. We included all the patients admitted to the Hospital Clínico Universitario of Santiago de Compostela, between January 2003 and December 2010 with a tentative diagnosis of ACS that was confirmed as myocardial infarction. There are 4,594 patients included in the database, 243 of them has an unknown residence or are missing data. The median length of follow-up for the patients was 1135 days (38 months). Among all the patients, 24% had died before the end of the follow-up period.

The response variable is considered as death for any cause before August 2011. There are many covariates included in the study such as components of the GRACE score (Granger et al. [3]) which is a measurement to assess the risk of ACS, spatial effects of municipalities of Galicia. The categorical covariates are: killip class (heart failure), ST-segment elevation myocardial infarction, in-hospital percutaneous coronary intervention. The continuous covariates are: age, heart rate, systolic blood pressure, creatinine, and troponine levels.

3. Geoadditive Regression Models and Predictive Accuracy.

The classical Cox regression models have some limitations in case of flexible modelling of geoadditive regression such as proportional hazard aasumptions, parametric form of the predictors and no spatial correlations. As an alternative analysis, we use structured hazard regression model, $\lambda_i(t) = exp(\eta_i(t)), i = 1,...,n$, with a geoadditive predictor;

$$\eta_i(t) = g_0(t) + X_i^{'}\gamma + \sum_{k=1}^{q} s_k(x_i k) + f_{(spat)}(s_i),$$

where $g_0(t) = log(\lambda_0(t))$ represents the log-baseline hazard rate, $s_k(x_k)$ is the nonlinear effect of a continuous covariate $x_k \cdot f_{(spat)}(s)$ is the spatial effect at site or in region s, and γ contains the usual linear effects.

The advantages of the structured hazard regression models are the possibility of joint estimation of baseline hazard function with the covariate effects, non-proportional assumptions of hazard function and including optional smoothing parameters. It also allows complex interactions and spatially correlated effects. This model is implemented in BayesX software freely available from www.bayesx.org. P-splines (Eliers and Marx [4]) were used to model continuous covariates and Markov Random Fields to model spatial effects (Kneib and Fahrmeir [2]).

The linear predictors of the model at time t are used to calculate the Incident sensitivity/Dynamic specifity to compute the ROC curves and the area under curve (AUC) for each time point t:

$$ROC_i^{(I/D)}(p) = TP'[FP_t^D](p);\ AUC(t) = \int_0^1 ROC_i^{(I/D)}(p)dp,$$