Bootstrap variable selection for tobit and logistic regressions to predict left ventricular contractility after acute myocardial infarction with a new panel of microRNAs

Olivier Collignon*[1], Yvan Devaux, [2] Melanie Vausort [2], Gerry P. McCann[3], Dominic Kelly [3], Leong L. Ng [3], Daniel R. Wagner, [2,4] Iain B. Squire [3]

[1] Competences Center for Methodology and Statistics, Public Research Centre – Health (CRP-Santé) , Luxembourg;

[2]Laboratory of Cardiovascular Research, Public Research Centre – Health (CRP-Santé), Luxembourg;

[3]Department of Cardiovascular Sciences, University of Leicester, and the NIHR Leicester Cardiovascular Biomedical Research Unit, Glenfield Hospital, Leicester, United Kingdom

[4] Division of Cardiology, Centre Hospitalier, Luxembourg.

*: olivier.collignon@crp-sante.lu

After acute myocardial infarction, left ventricular remodeling occurs frequently in patients, which induces a loss of cardiac function. Wall Motion Index Score (WMIS) is widely used as a surrogate of left ventricular contractility impairment after acute myocardial infarction. Predicting such a clinical outcome would be useful but challenging. In this study, in 150 patients with acute myocardial infarction, we evaluated the predictive power of a new panel of 4 microRNAs (miRNAs) added to a clinical and biochemical reference model. All 15 combinations of miRNAs were generated and successively added to the reference model in order to assess the best improvement in prediction possible. Since WMIS is a left censored continuous variable, tobit regression was performed for each new list of predictors and Akaike Information Criterion (AIC) was used as the optimality criterion to select the best fit. Outcome prediction was evaluated using Mc Fadden's R^2. WMIS was also treated as a binary outcome (WMIS <= 1.2, WMIS >1.2) as is often done in medical studies. In the same way, the best model was selected via AIC minimization in logistic regression. WMIS prediction was then assessed by continuous Net Reclassification Index (NRI), Integrated Discrimination Improvement (IDI). In order to avoid over-fitting and to correct statistics for optimism, the whole scheme was validated internally on 150 bootstrap samples. For logistic (resp. tobit) regression, the best combination of biomarkers added was the 4 miRNAs, p=0.005 (resp. 0.049), (Likelihood Ratio test compared to the reference model) since it was the most often selected combination in bootstrap samples (59% (resp. 29%)). For logistic regression, the optimism corrected statistics were: continuous NRI=0.504 (p=5.05E-05) and IDI=0.048 (p=0.001), whereas for tobit regression we obtained Mc Fadden's R^2=0.09. In conclusion, this procedure was useful to evaluate the predictive ability of this new panel of miRNAs. Further experiments and new datasets are necessary to validate these findings.

Key Words: Modeling, Wall Motion Index, biomarker, heart attack