ABSTRACTS - 28th EACTS

Prediction: Art or science
Monday, 13 October 2014

024
IMPROVED PREDICTION BY DYNAMIC MODELLING: AN EXPLORATORY STUDY IN THE ADULT CARDIAC SURGERY DATABASE OF THE NETHERLANDS ASSOCIATION FOR CARDIO-THORACIC SURGERY


1Department of Cardiothoracic Surgery, LUMC, Leiden, Netherlands; 2Department of Public Health, Erasmus University Medical Center, Rotterdam, Netherlands; 3Thoracic & Cardiac Surgery, University Hospital St Radboud, Nijmegen, Netherlands; 4Department of Cardio-Thoracic Surgery, VU Medical Center, Amsterdam, Netherlands; 5Department of Cardiothoracic Surgery, Erasmus MC, Rotterdam, Netherlands

Objectives: The predictive performance of risk prediction models may deteriorate in time. We aimed to explore different methods for continuous updating of models ('dynamic modelling') in order to improve risk prediction.

Methods: Data on adult cardiac surgery from 2007 to 2012 (n = 95,240) were extracted from the Netherlands Association for Cardio-Thoracic Surgery database. The logistic EuroSCORE predicting in-hospital death was updated using five methods: recalibrating the intercept of the logistic regression model, recalibrating the intercept and joint effects of the risk factors, re-estimating all risk factor effects, closed test procedure to select either of these three and a Bayesian learning strategy. Models were updated with 1 or 3 years of data, in all cardiac surgery or within operation subgroups. Performance was tested in the subsequent year using discrimination (area under the ROC-curve, AUC) and calibration (calibration slope and calibration-in-the-large).

Results: Compared to the EuroSCORE, all updating methods resulted in improved calibration-in-the-large (–0.17–0.04 vs –1.13–0.97, ideally 0.0), calibration slope (0.92–1.15) and discrimination (AUC 0.83–0.87) were comparable. Results were similar in the subgroup isolated coronary artery bypass graft (CABG). In the smaller subgroups aortic valve replacement (AVR) and AVR + CABG, extensive updating (re-estimation all risk factor effects) using 1 year of data led to poor performance.

Conclusion: Several methods for dynamic modelling resulted in good discrimination and superior calibration compared to the EuroSCORE. The preferred method depends on the size of the database: extensive modelling such as re-estimation of all risk factor effects requires more data than recalibration. In smaller databases or subgroups, using data from multiple years or the Bayesian approach are recommended.