Predicting prognosis in cardiac surgery: a prophecy?

A. Pieter Kappetein* and Stuart J. Head

Department of Cardio-Thoracic Surgery, Erasmus University Medical Center, Rotterdam, The Netherlands

* Corresponding author. Department of Cardio-Thoracic Surgery, Erasmus University Medical Center, PO Box 2040, 3000 CA Rotterdam, The Netherlands. Tel: +31-10-4635412; fax: +31-10-4633993; e-mail: a.kappetein@erasmusmc.nl (A.P. Kappetein).

Keywords: EuroSCORE • Risk reduction • Cardiac surgery

Physicians need to make predictions on the prognosis of a treatment that helps them in the choice of therapy. Medicine used to be much more subjective than in the current evidence-based era. Shared decision-making, where physicians and patients both participate in deciding on choices for therapy, is also more common [1]. Clinical prediction models like EuroSCORE may provide the evidence-based input for shared decision-making by providing an estimate of the operative risk of patients undergoing cardiac surgery. An ideal clinical model would be something simple dividing the patients into ‘good’ and ‘bad’ without further specification of the survival chances. The original EuroSCORE was a compromise between the ‘statistical ideal’ and the ‘clinical ideal’. It was developed from a large multinational European population and was a model predicting mortality based on 17 variables, either from a logistic regression equation or from an additive model. Numerous institutions throughout the world have tested and validated EuroSCORE.

Loss of calibration with the additive and logistic EuroSCORE has been observed by many investigators and an update of the EuroSCORE was warranted. One possible reason for the poor calibration of the original EuroSCORE score is that the score was developed from patients undergoing surgery almost 20 years ago. As surgical and perioperative care evolves and the impact of clinical variables change, prediction models therefore require revision. These factors may also vary between institutions and it is well known that the quality of care and comorbidities of patients differ between countries. The original EuroSCORE already identified major differences in the risk profile of national samples [2]. This is therefore one of the major concerns with EuroSCORE II: 154 hospitals from 43 countries participated, of which many were outside Europe [3]. One may, therefore, question whether the term EuroSCORE is still valid or another name should be used that reflects the fact that so many countries outside Europe participated. With this in mind, it becomes even more important that, as indicated by the authors, units and surgeons calculate their own risk-adjusted mortality ratio. The model is probably more reliable in the prediction of death over a wide range of risk groups rather than the prediction of the vital status of an individual patient.

Another reason for the poor calibration in the original EuroSCORE might be that a large number of risk factors in the model are highly correlated. It is important to recognize correlation between predicting variables, as the additional risk contribution of certain variables can in some part be explained by the effect of other variables. Some predicting variables may also be more important for some types of operations then for others. The large number of risk factors with potential interaction may overestimate risk in certain categories of patients (e.g. intermediate risk or extreme risk). It is therefore a pity that the authors have not explored possible interaction terms in the new EuroSCORE II, something the Society of Thoracic Surgeons score has taken into account.

For the analysis, the authors chose to drop cases with missing data. Besides inefficient use of available data, bias may arise due to systematic differences between subjects with complete data and subjects with missing data. An estimated regression coefficient for a predictor might be influenced if the missing data are associated in some way with the outcome [4]. They could have chosen to use some form of imputation to preserve those cases. In general, the quality of those centres with missing data or those unable to provide specific outcome data may be questioned.

One of the major concerns with EuroSCORE II is that the primary outcome was mortality at the base hospital. In current practice, however, it is common that patients are transferred to referring hospitals at different points in time after the operation. There is significant geographic and hospital variance with regard to the day of transfer. For example, the length of stay in coronary bypass patients in the SYNTAX trial ranged from a mean of 5–20 days. A fixed point in time in a mortality prediction model has advantages over the current model, as it provides the ability to compare centres. Current guidelines and clinical trial practices mandate mortality assessment at 30, 60 or 90 days [5, 6]. The number of centres that provided 30- or 90-day mortality was disappointing.

The authors are to be admired for the amount of work they have put into the new model and for their energy in starting already on a EuroSCORE III project. We have to be careful, however, not to add prognostic factors all the time. Models with only a few parameters are quite stable and estimating a few calibration parameters might be enough [7]. ‘Garbage in, garbage out’ is a well-known problem inherent to risk models, causing inaccurate risk prediction. The inclusion of a greater number of variables increases the risk of errors that can be caused by differences in the interpretation of definitions, typing errors or conflicting chart information [8].

It took many years to learn the advantages and shortcomings of EuroSCORE I. Many institutions have adopted the EuroSCORE in their quality control programmes. Implementing EuroSCORE II and learning the benefits will also take some time. There is
Currently more need for models that not only focus on mortality but also on postoperative complications and the development of procedure-specific models. As clinicians are confronted with more elderly patients, it might also be useful to focus on specific subsets of patients. A prognostic model is only useful if its predictions are at least as accurate as those of the doctors who would use it. We have to be thankful to the initiators of the EuroSCORE project for their great contribution to our profession.

Conflict of interest: none declared.

REFERENCES


