The New EuroSCORE Project

Samer A.M. Nashef

on behalf of the EuroSCORE Project team

Readers of this journal will be familiar with the European System for Cardiac Operative Risk Evaluation (EuroSCORE) as the world's most widely used cardiac surgical risk model. The model was built from data of patients operated on in 1995 and first published in 1999, so it is now ten years old.

Since the introduction of EuroSCORE, there has been a quantum improvement in cardiac surgical survival which occurred in the first two to three years of the new millennium. Evidence from countries with national databases suggests that mortality in some of them has approximately halved, despite gradual worsening of the risk profile of patients. In the United Kingdom, for example, mortality has fallen to approximately 55% of logistic EuroSCORE prediction. This phenomenal improvement in cardiac surgical outcomes may be largely due to the Hawthorne effect: this is the well recognised improvement in performance which occurs simply when performance is measured. Until the widespread use of EuroSCORE, there was no established measure of cardiac surgeons' clinical performance. EuroSCORE provided the tool for such measurement, and performance improved.

The test of a risk model is in assessing two features: calibration and discrimination. Calibration is the accuracy of the model for predicting risk in a group of patients, in other words, if the model says that mortality in a thousand patients is likely to be 5%, and actual mortality is 5% or close to it, then the model is well calibrated. Discrimination refers to the model's ability to distinguish between low-risk and high risk patients. In other words, if most of the deaths occur in patients that the model identifies as high risk, the model has good discrimination, but most deaths occur in patients that the model identifies as low risk, there is poor discrimination. We measure discrimination using a statistic called the "area under the receiver operating characteristic (ROC) curve". If the area under the ROC curve is 0.5, the model does not discriminate at all. Good discrimination begins at 0.7 and rarely exceeds 0.85. If the area is 1.0, the model is no longer a risk model but a crystal ball which forecasts the future (an impossible task).

EuroSCORE remains powerful in discriminating between low-risk and high-risk patients, and to this day the area under the ROC curve in many studies is around 0.8, but its calibration has been called into question by many centres, stating that it overpredicts risk (a few centres still say it underpredicts risk). The time has therefore come to find out if the model should be recalibrated. We shall be collecting new data for this and will take the opportunity to refine and improve the risk factors and their assessment so as to improve further not just calibration, but also discrimination.

We could do this in two ways. One way is to use recent data from the many databases of cardiac surgery that exist throughout the world. Another is to collect new data. Most databases are unfortunately not validated and we are not able to verify their quality. One reason for the success of EuroSCORE is that it was built on a robust and "clean" database, collected by volunteer centres. We would like to keep that standard, and will therefore collect new data.

This time, participation will be easier. In 1995, we asked for around a hundred data points on each patient. The original risk model has already discarded most of these factors as unhelpful in risk assessment. Conversely, evidence from studies suggest that new risk factors should be added and old ones refined. This time, we shall ask for only around 20 data points, similar to data already collected by anyone who uses EuroSCORE, with a few additions and refinements.

These will be in the areas of diabetes, obesity, renal function, the weight of the proposed intervention and one or two other fields. Data collection will be online via a dedicated website. There will be an option to collect paper-based data for later online transfer for centres requiring that facility. Data will be confidential, guaranteeing patient, centre and surgeon anonymity to all participants.

It is crucial that the data come from all kinds of centre, and not only so-called "centres of excellence". The entire spectrum of performance must be included and all deaths must be reported so that the risk model reflects the reality of cardiac surgical outcomes. In fact, if only the centres with the best outcomes participate, the risk model will be harsh on everyone else.

The website is under construction and will be both robust and user-friendly. Data collection will begin early in 2010. The exact time period for collecting data is not yet fixed, because that will depend on the number of centres participating. The more centres participate, the quicker the data collection will be completed. Already more than 100 centres have committed to the project. We hope that number will increase very rapidly. If enough centres participate, we may have data on 40,000 patients in less than 3 weeks. The longest we shall ask for will be 3 months.

Centres in Poland with an interest in participating in this important inititative are invited to register their interest by

email (euroscore@papworth.nhs.uk) or by following the link from the EuroSCORE website (www.euroscore.org).

The EuroSCORE Project Team is the same scientific team behind the original EuroSCORE. We remain independent of national and international specialist societies, governments and industry. The project is currently funded by a scientific grant from Edwards Laboratories, by Papworth Hospital, Cambridge and by the University of Stockholm in Sweden. Additional funding sources are also being sought. Our only aim is to produce the finest and most practical risk model for cardiac surgeons and their patients.

In the meantime, users of EuroSCORE can be assured that it is still a valuable tool for assessing cardiac surgical risk. Any risk model offers a set standard. Some units will perform at that standard, some will do better and some worse. The best estimate for evaluating the risk of mortality for a patient undergoing a particular procedure at a particular institution is to use logistic EuroSCORE and then correct it for the performance of the unit in question, so that the patient should be quoted a predicted mortality calculated as follows:

predicted mortality = patient logistic EuroSCORE × hospital mortality

hospital logistic EuroSCORE

in other words, in Hospital X, where actual mortality is, say, 0.7 of predicted, the patient's logistic EuroSCORE is multiplied by 0.7, reflecting the hospital's performance.