The challenge of departmental quality control in the reengineering towards off-pump coronary artery bypass grafting

Paul Sergeant*, Erik de Worm, Bart Meyns, Patrick Wouters
Cardiac Surgery Department, Gasthuisberg University Hospital, 3000 Leuven, Belgium
Received 5 October 2000; received in revised form 6 April 2001; accepted 6 June 2001

Abstract

Objective: Off pump coronary surgery is a major reengineering effort of the surgical systems. There are no perfect tools available to guide every centre in the confrontation with the complete spectrum of risk and the limited number of events. This study analyses the use of a hospital mortality risk-stratifying system in the complete shift towards off-pump CABG.

Methods: All 535 off-pump CABG patients from January 1997 till September 2000 underwent a comparison of their hospital mortality versus the EuroSCORE predictions. The mean risk predicted by the EuroSCORE was 4.5 ± 3% (range 0–14) and the mean age was 65 ± 10 years (range 36–89). The series includes 23 repeat procedures, also 77 patients with per oral or insulin-treated diabetes. The number of distal anastomoses was 2.5 ± 1 and of arterial grafts 1.3 ± 0.6.

Results: The observed hospital mortality was 15 patients, 2.8% (Fisher exact test P > 0.19 versus the EuroSCORE). The 1 and 3 month Kaplan–Meier survival, irrespective from hospital discharge, was 97.4 ± 0.7 and 97.2 ± 0.7%, respectively. A cumulative risk-adjusted mortality plot is constructed. The area under the ROC curve was 0.886. A stepwise sampling of patients according to increasing risk identified the difference between the EuroSCORE-predicted and observed hospital mortality for the complete spectrum of risk. The P value of this difference was 0.06 for the grouping including all patients from 0–5% risk (78% reduction), 0.04 for the grouping 0–8% risk (61% reduction), and 0.05 for the grouping 0–11% risk (52% reduction of risk). The loss of statistical significant difference was due to the inclusion of the patients at extremely high risk.

Conclusion: A hospital mortality risk-stratifying system can provide guidance but different and in depth approaches are mandatory to improve the insight, certainly in the presence of a large spectrum of risk.

Keywords: Coronary Artery Surgery; Off-pump CABG; EuroSCORE

1. Introduction

Coronary bypass surgery is one of the most commonly performed [1] surgical interventions, one of the more expensive ones [2] and nevertheless associated with peri-procedural mortality and morbidity (http://www.sts.org/doc/4246). Attending physicians, surgeons and anaesthesiologists, have the permanent societal commitment to reengineer the procedural hospital stay towards a reduction of mortality, morbidity and costs. The reduction or complete avoidance of extra-corporeal technology could be one of the tools to obtain this reduction.

Off-pump coronary surgery is in most centres restricted to stable patients with only anterior or proximal inferior coronary disease. The hospital mortality using extra-corporeal technology in these patients has always been very low [3], so a considerable cohort of several thousands of patients would be needed to identify a benefit in randomised trials. We decided to expand the patient-spectrum into the high-risk domain to increase the likelihood of a possible benefit and to monitor closely the observed result.

The purpose of this manuscript is to study how the most recently published European multi-institutional risk-stratifying system, the EuroSCORE [4], could help a centre in the evaluation of the re-engineering towards off-pump CABG.

2. Materials and methods

2.1. Patient material

This dataset includes all 535 CABG procedures started and finished without the use of extra-corporeal circulation at the Katholieke Universiteit of Leuven (Belgium) from January 1997 till September 2000. Patient selection for OPCAB
procedure was most frequently based on low risk and anterior or proximal inferior coronary disease from January 1997 till September 1999, these were the first 118 patients. The learning curve of this surgical technique is therefore included in this manuscript. From October 1999 the approach became the procedure of choice, including 90% of all isolated primary and repeat CABG at the Department. The surgical technique in the other patients was a rotary support system or cardio-pulmonary bypass for patients in a randomised trial and cardio-pulmonary bypass for patients in cardiogenic shock or cardio-pulmonary resuscitation before surgery.

There were 414 male and 121 female patients. The mean age of the patients was 65 ± 10 years with a range from 36 to 89 years, 80 patients were older than 75 years. Single vessel disease was present in 9%, two vessel disease in 32% and triple vessel disease in 59% of the patients. Diabetes was treated with oral medication in 53 patients and with insulin in 24 patients. The CABG was a repeat procedure in 23 patients. The average EuroSCORE risk was 4.5 ± 3% (see Table 1 for the number of patients at risk in every risk category).

Conversion after an initial off-pump approach into regular extra-corpooreal technology was only needed for technical reasons in four patients, none of these patients died during hospital stay, these patients were not included.

2.2. Surgical technique

Patients were premedicated with 0.05 mg.kg⁻¹lorazepam sublingually 30 min prior to admission to the operating theatre. A large bore IV line and a 20 gauge radial artery catheter were placed and anesthesia was induced intravenously with midazolam 0.05 mg.kg⁻¹, pancuronium 0.1 mg.kg⁻¹ and a target controlled infusion (TCI) of propofol at 1 mg.ml⁻¹. After orotracheal intubation and the start of mechanical ventilation, a thermodilution catheter and central venous line were installed. Maintenance of anesthesia consisted of propofol TCI and intermittent bolus doses of sufentanil. Two-dimensional transesophageal echocardiography was used selectively in patients with pre-existing left ventricular dysfunction, suspected atheromatosis of the ascending aorta or valvular lesions. Continuous multiplane ST-recording monitored myocardial ischaemia. Throughout the procedure, heart rate, arterial systolic blood pressure and cardiac index were maintained between 60 and 90 beats/min, 85–120 mmHg and 2–3 l.min⁻¹m⁻² respectively using atrial pacing, beta-blockade, neosynephrine and optimisation of preload with colloid fluids and legs-up position where necessary. The use of beta-1 agonists was avoided to all possible extent.

A midline sternotomy was the standard approach for all cases in this dataset. The left thoracic cavity was usually opened for the prelevation of the left internal mammary graft. The pericardium was then suspended to the left side. The right thoracic cavity was only opened in case of bilateral internal mammary artery grafting. The right-sided pericardium was never suspended during the distal anastomoses time, allowing mobility of the heart towards the right chest. A V-shaped sponge was fixed down on the right posterior mediastinum halfway in the line joining the inferior cava to the left inferior pulmonary vein. Closing and retracting the V shape, enucleated the left atrium and similarly the left ventricle. The anastomotic zone was further stabilised with compression- or suction devices. The coronary vessels were shunted where possible, without causing endothelial damage.

Myocardial protection was guaranteed with Lidoflazine intravenous at 1 mg.kg⁻¹, administered over 30 min. The surgical strategy towards complete revascularisation and arterial grafting was unchanged versus the on-pump practice. The number of distal anastomoses was 2.5 ± 1 (range

Table 1

<table>
<thead>
<tr>
<th>EuroSCORE predicted value</th>
<th>N at risk</th>
<th>N of predicted deaths</th>
<th>N of observed deaths</th>
<th>% observed mortality</th>
<th>95% lower confidence limits of observed mortality</th>
<th>95% upper confidence limits of observed mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>55</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.5</td>
</tr>
<tr>
<td>1</td>
<td>43</td>
<td>0.43</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>8.2</td>
</tr>
<tr>
<td>2</td>
<td>60</td>
<td>1.2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6.0</td>
</tr>
<tr>
<td>3</td>
<td>66</td>
<td>1.98</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5.4</td>
</tr>
<tr>
<td>4</td>
<td>59</td>
<td>2.36</td>
<td>1</td>
<td>1.7</td>
<td>1.6</td>
<td>7.4</td>
</tr>
<tr>
<td>5</td>
<td>60</td>
<td>3</td>
<td>1</td>
<td>1.7</td>
<td>1.6</td>
<td>7.2</td>
</tr>
<tr>
<td>6</td>
<td>52</td>
<td>3.12</td>
<td>3</td>
<td>5.8</td>
<td>4.6</td>
<td>10.1</td>
</tr>
<tr>
<td>7</td>
<td>47</td>
<td>3.29</td>
<td>1</td>
<td>2.1</td>
<td>2.1</td>
<td>9.2</td>
</tr>
<tr>
<td>8</td>
<td>33</td>
<td>2.64</td>
<td>1</td>
<td>3.0</td>
<td>2.9</td>
<td>12.7</td>
</tr>
<tr>
<td>9</td>
<td>21</td>
<td>1.89</td>
<td>3</td>
<td>14.3</td>
<td>11.3</td>
<td>22.0</td>
</tr>
<tr>
<td>10</td>
<td>17</td>
<td>1.7</td>
<td>1</td>
<td>5.9</td>
<td>5.7</td>
<td>22.8</td>
</tr>
<tr>
<td>11</td>
<td>10</td>
<td>1.1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30.8</td>
</tr>
<tr>
<td>12</td>
<td>6</td>
<td>0.72</td>
<td>2</td>
<td>33.3</td>
<td>29</td>
<td>44.4</td>
</tr>
<tr>
<td>13</td>
<td>3</td>
<td>0.39</td>
<td>1</td>
<td>33.3</td>
<td>32.5</td>
<td>57.2</td>
</tr>
<tr>
<td>14</td>
<td>3</td>
<td>0.42</td>
<td>1</td>
<td>33.3</td>
<td>32.5</td>
<td>57.2</td>
</tr>
</tbody>
</table>
1–5) and the number of grafts 2.3 ± 0.7 (range 1–4). Unilateral and bilateral mammary artery grafting were performed in 294 and 204 patients, respectively, resulting in an average of 1.3 ± 0.6 arterial grafts per patient.

2.3. Statistical methods

The studied event is the hospital mortality, irrelevant of the time interval between surgery and discharge. Continued hospital stay in a hospital different from the surgical treating centre is included in this analysis. The patients were entered into the study till September 1st 2000 and data closure was September 15th 2000. All patients still hospitalised at that date were considered hospital survivors.

All patients were contacted by the authors at the 3-month anniversary of their date of surgery. Kaplan–Meier survival results were constructed irrespective from hospital discharge.

Continuous data were presented as mean ± standard deviation.

The analysis starts with an overall evaluation between the EuroSCORE risk-adjusted predicted and the observed hospital mortality using the Fisher exact test.

The CRAM [5,6] plot (Fig. 1) depicts the cumulative EuroSCORE risk-adjusted mortality. This is a cumulative plot depicting the lives saved or lost versus the EuroSCORE, adding value if a patient survives the hospital stay (e.g. +0.05 if a patient with 5% predicted risk survives) and subtracting value if a patient dies during hospital stay (e.g. − 0.95 if a patient with 5% predicted risk dies). If the line equals the baseline then the observed risk equals the predicted risk, if the line moves above the baseline then the observed result expresses the number of lives gained versus the EuroSCORE.

A receiver operating characteristic curve (ROC) [7,8] evaluated the predictive performance (discriminatory power) of the EuroSCORE. This was created by using each EuroSCORE value (0–14) as a theoretical cut off point to predict in-hospital mortality. The sensitivity and the specificity of the prediction were calculated for each EuroSCORE value. The sensitivity was then plotted versus the 100-specificity and the points were interconnected. The area under this curve is a measure of the discriminatory power of the test.

A virtual dataset is constructed with the same number of patients at risk for every risk category with the EuroSCORE risk as predicted outcome. The analysis continues (Table 1) with the comparison for each risk category (from 0 to 14%) of the predicted versus the observed hospital mortality. The uncertainty of the observations due to the limited number of patients at risk for every risk category is corrected by presenting also the 95% upper and lower confidence limits of the observed mortality.

The analysis finishes with a stepwise approach towards overall comparison (Table 2, Figs. 2 and 3). Each column in the figures or each line in the table is a stepwise increasing risk sample. For each sample a Fisher exact test is performed.

![Fig. 1. The cumulative EuroSCORE risk-adjusted mortality of the off-pump CABG patients (N = 535).](image-url)
between the predicted and the observed hospital mortality. The data were analysed as if the EuroSCORE virtual sample was an independent group. In the first column or line only patients with 0% predicted risk are included, in the second column only patients with 0 and 1 risk are included. The last column is the complete study sample. The % difference is calculated using the formula: \( \frac{(\text{predicted deaths}) - (\text{observed deaths})}{\text{predicted deaths}} \times 100 \). This method simulates the expansion of the off-pump program into the higher risk domain.

3. Results

The in-hospital observed mortality was 15 patients (2.8%). There was no significant \( P = 0.19 \) difference between the overall predicted (4.5%) and observed mortality. Fig. 1 depicts the CRAM (cumulative EuroSCORE risk-adjusted mortality) plot of the study population. The average hospital stay of the hospital survivors was 11 ± 8 days (median 9).

The 1, 2 and 3-month Kaplan–Meier survival, irrespective from hospital discharge, was 97.4 ± 0.7, 97.2 ± 0.7 and 97.2 ± 0.7%, respectively.

The area under the ROC curve of the EuroSCORE versus the observed in-hospital mortality was 0.886 (95% confidence interval 0.85–0.91).

The validation by risk or EuroSCORE category (Table 1) identified predicted mortalities laying between the 95% upper and lower confidence limits of the observed results for all categories. The observed results for categories 0–8 and 10 and 11 were individually lower than the predicted result.

The stepwise cumulative approach (Table 2, Figs. 2 and 3) identified no observed mortality for the category of patients where no hospital mortality was predicted (Fisher exact test \( P = 1 \)). Increasing stepwise the risk and thereby the sample size, by including all patients up to 3% predicted hospital stay \( (N = 224) \) identified no observed mortality for 3.6 predicted deaths (Fisher exact test \( P = 0.12 \)).

Further increasing the risk and sample size to 475 patients, including all risk up to 8%, identified a difference between observed and predicted mortality of 61% (Fisher exact test \( P = 0.04 \)). This difference remained significant \( (P = 0.05) \) up to the 0–11% risk sample and decreases beyond that sample size.

4. Discussion

The analysis of the incremental benefit of a change in surgical procedure in patients with a large risk spectrum is a very complex statistical problem. There are a number of methodological options. The most optimal option is the randomised trial. The trial cannot be blinded for obvious reasons and the alternative therapy has a variety of technical possibilities. The spectrum starts with the cardioplegic arrest and core cooled approach and finishes with the normothermic non-vented beating supported heart. Therefore all possible inferences from a randomised trial will stay inconclusive as long as not all alternatives will have been studied. The variability in cardiac and non-cardiac comorbidity mandates certainly a multivariate correction. Each correcting variable (whether a scoring value or an independent variable) requires a number of events and will therefore increase the sample size required. Off-pump coronary surgery is usually started in patients with low risk, requiring an additional increase in sample size. These sample sizes can only be found in large multi-centre datasets, adding possible strata to the data. A randomised trial for a specific subset might become interesting after the creation of inferences from observational studies.

The second methodological option is the comparison with a simultaneous same-site experience using an alternative surgical approach. The alternative dataset can be expanded.
in the more recent years to include enough patients but the study dataset will similarly require enough patients to allow correction for patient-variability. Even after extensive correction a legitimate doubt of patient-selection will remain. With a total mortality of only 15 patients in the total off-pump experience no such approach and multivariate correction was possible. In addition, the quality of care in the alternative dataset would always remain a standard open to criticism.

We have opted for the third methodological option by using the EuroSCORE for different methods of analysis, thereby solving some of the previous issues. The EuroSCORE scoring system is a recently published system based on a large \(N = 13302\) patients multinational \(N = 8\) countries, multi-centre \(N = 132\) centres dataset [9]. This scoring system predicts only the hospital mortality. This is definitely a biased segment of the peri-procedural risk [9] but the risk of hospital death decreases very rapidly a few days after surgery. We have tried to circumvent this limitation by including all mortality even in a secondary referral centre. The longest hospital stay of 120 days in one patient and the median hospital stay of 9 days should cover most of the expected and unexpected death. The Kaplan–Meier results did not identify any additional risk between hospital discharge and the end of the third month after the date of surgery. We have tried to circumvent this limitation by including all mortality even in a secondary referral centre. The longest hospital stay of 120 days in one patient and the median hospital stay of 9 days should cover most of the expected and unexpected death. The Kaplan–Meier results did not identify any additional risk between hospital discharge and the end of the third month after the date of surgery. Hence the hospital mortality analysis made sense as an indicator for the quality of care.

The risk profile of the study sample \((4.5 \pm 3\% )\) exceeds the risk profile of the EuroSCORE dataset \([11]\) \((3.3 \pm 2\% )\) as well in average value as in standard deviation, indicating the representativeness of the study sample for the current coronary surgery population in Europe. The number of distal anastomoses is similar to the number of anastomoses in the Italian, French and Spanish subsets. The practice of arterial grafting exceeds by far the current practice of an average of 0.7–1.1 arterial grafts per patient in Europe according to the same EuroSCORE dataset. The initial off-pump experience, started in 1997, was included in this dataset.

The first crude observation of the in-hospital mortality \((2.8\% )\) was inconclusive and falls within an expected mortality of an unspecified dataset of CABG patients (http://www.sts.org/doc/4246). The second approach by comparing the overall predicted versus the observed mortality indicated a 38\% reduction of risk but remained similarly inconclusive. The CRAM plot depicts this positive behaviour but the methodology does not include confidence intervals, so no additional inferences can be build. The Kaplan–Meier survival curve limits some of the uncertainty by identifying the absence of early risk beyond hospital discharge.

The area under the ROC was considerable, but two limitations of this observation prevail. This measure indicates that the EuroSCORE has considerable discriminatory power in the risk-adjustment of the in-hospital mortality after off-pump surgery but gives no information about the accuracy of the prediction. In addition a sample size of 50 events is often suggested as a minimum for accurate analysis.

The validation by EuroSCORE risk category is obviously limited by the limited number of patients in each risk category and is once more inconclusive. Since most centres would probably explore the high risk CABG domain in a stepwise fashion we tried a similar stepwise approach in our analysis to explore evidence of additional cost or benefit. The off-pump experience demonstrated a superior performance in hospital mortality versus the EuroSCORE standard, this difference was considerable but far from significant if only low or average risk patients were included in the dataset and reached only borderline statistical significance for the 0–8 and 0–11 samples. The gradual decrease beyond the 11\% risk pivot of the difference between observed and predicted hospital mortality leads to the loss of statistical significance and could be caused by several conditions. It is possible that the highest risk patients do not benefit from avoiding the extra-corporeal technology. This is obviously true for the patients brought to surgery in cardio-pulmonary resuscitation. The loss of pulmonary function through pulmonary oedema mandates also extra-corpooreal technology in the surgical approach of patients in cardiogenic shock. These last two categories were not included in this off-pump dataset.

The second condition is the limited number of patients \((N = 12)\) at risk in this study sample beyond the 11\% pivot, thereby creating considerable uncertainty. The third condition is the possible underscore of high risk by the EuroSCORE scoring system. The EuroSCORE publications give no information about the spectrum of risk included in their core datasets.

In conclusion the use of some risk-adjusting system, as the EuroSCORE, is essential in the institutional search for off-pump surgery benefit. Different approaches are simultaneously necessary to evaluate this benefit. The inclusion of only low risk patients in the practice of off-pump surgery will exclude the possibility of the analysis, unless a dramatic increase of mortality is observed. A dilemma is therefore created. A possible risk-reducing effect will in most centres only become visible by including considerable risk. Extensive experience in lower risk patients, difficult to evaluate, will have to preclude this possible finding. More than 500 patients were needed in this study population. The present publication presents a methodology of real-time monitoring, essential in this extreme reengineering of the coronary bypass procedure.

References

forward for coronary revascularization. We surgeons seem preoccupied with mortality. What would be interesting is the issue of ‘near miss’, that is, the issue of morbidity falling short of death. It would be nice to hear about length of stay, it would be nice to hear about vessel occlusion, about perioperative myocardial infarction and so on. Do you have any data beyond the EuroSCORE?

Dr Sergeant: Of course we have these data. The point is that we wanted to have a very focused study on a very precise issue. If we want to convince cardiologists and patients, the very first issue that comes up to mind is: are there lives saved?

Dr M. Turina (Zurich, Switzerland): Just adding a word of caution and trying to hear your opinion, all the studies with the off-pump usually have less anastomosis than on-pump, and this is a common trait to all previously published studies about MIDCAB. I haven’t quite caught if there is a difference in the number of anastomoses you are doing now, but it is obvious that some of these posteriorly located anastomoses might be difficult.

My comment is that probably it is not enough to look only at the operative mortality, because some of the problems due to the less complete revascularization will be visible only a year or two later, and this is what also makes the basic difference between PTCA and stenting vs. coronary artery bypass grafting. So I would be very interested to hear your data a year from now and see if there might be some difference due to the different strategy of revascularization.

Dr Sergeant: The point is well taken. There is a minimal difference of the number of anastomoses, like 10%, versus what is available within the EuroSCORE data set, but the difference is due to the first 118 patients, as I have already said, who were selected patients. If one looks at the last 400 patients, then the total number of anastomoses in our center, which we are monitoring since 1971, has absolutely not changed. And with the current techniques of enucleation, as we call it, and stabilization, the number of anastomoses on the lateral or posterolateral wall is absolutely unlimited.

There is no reason why you wouldn’t be able to do four or five anastomoses if that would be your challenge to do on the back side of the heart.

You are absolutely right about the issue of the long-term information, but if it is technically not possible to have this information. We have added that one of the first limitations of studying hospital mortality is that it is only a part of the early risk, and that is why we have included in the manuscript the 3-month complete actuarial survival using the Kaplan–Meier method. There is no additional mortality in the first 3-month interval beyond hospital discharge.

Dr J. Svennevig (Oslo, Norway): I also would like to ask you, is it your center that outperforms the EuroSCORE? That could be answered probably by asking you, did you apply other scoring systems to your observed mortality, and would it be possible to repeat this analysis in another center?

Dr Sergeant: Yes. The methodology of analysis is a very simple one. It’s running with Excel with now and then a little sidestep towards Statview. But it can be done in any office. In fact, I am suggesting to colleagues that this is a very simple method of measuring their performance in off-pump surgery, and whether it is just a center or the technique is something which others will have to prove.

Dr Svennevig: I think there are a lot of European centers that are very interested now in the EuroSCORE, a lot of centers that have not been using scoring systems so far and that are really considering the EuroSCORE. So it would be very interesting to know the results.

Appendix A. Conference discussion 34

Mr S.A.M. Nashef (Cambridge, UK): This was very interesting, and it reminded me of a quote that one of our statisticians used, that if the results at first don’t suit your desired outcome, it is time to manipulate the data. I think there are two possible conclusions that can be drawn from your paper. The first one, of course, is that OPCAB outperforms the EuroSCORE, the second one, of course, is that your center outperforms the EuroSCORE.

Dr Sergeant: This study is an exploration and not a manipulation of data. The issue of whether this is a center issue or a procedure issue can only be solved in a stepwise manner. Including only low risk patients in the patient material will never generate knowledge or identify benefit. It’s only by providing this procedure for the whole spectrum of patients that one can start making inferences. I would really invite other centers to look at their data in the same way.

Mr S. Large (Cambridge, UK): This is only 500 patients with very small numbers, and I would like to just back up what Mr. Nashef said in his criticism.

I would like to try and convince myself that OPCAB is the right way forward for coronary revascularization. We surgeons seem preoccupied with mortality.