American University of Armenia
Center for Health Services Research and Development

Nork Marash Medical Center

COMPARISON OF THREE MODELS TO PREDICT OPERATIVE MORTALITY RISK FOR CORONARY ARTERY SURGERY IN YEREVAN NORK MARASH MEDICAL CENTER

Lusine Abrahamyan, MD, MPH
Anahit Demirchyan, MD, MPH
Michael E. Thompson, MS, DrPH

Yerevan, 2004

TABLE OF CONTENTS

Abstract

Objective- To compare the ability of three different risk models to predict 30-day operative mortality after isolated coronary artery bypass graft surgery in Yerevan Nork Marash Medical Center (NMMC).

Methods- Retrospective medical record review. Inclusion criteria: all patients who underwent isolated CABG surgery in 2003. Risk factors were recorded for all patients, along with operative (30-day postoperative) mortality. Predicted mortality was determined from EuroSCORE, UK Bayes Simple & UK Bayes Complex models.

Results- The observed mortality was 2.1% (8 of 379). The mean predicted mortality by EuroSCORE, UK Bayes Simple and UK Bayes Complex model was 2.36%, 1.53% and 1.46% respectively. There were no statistically significant differences between the observed mortality and predicted mortality rates from any of the models ($p>0.05$). The overall predictive ability of EuroSCORE, UK Bayes Simple and UK Bayes Complex models in terms of their discrimination as measured by the area under the Receiver Operating Characteristic (ROC) curve were 0.84, 0.869 and 0.864 respectively. There were not significant differences among the areas under the ROC curves. Model calibration was measured by Hosmer-Lemeshow goodness-of-fit test. The resulting *p* values of chi-square statistics were 0.71, 0.10 and 0.38 for EuroSCORE, UK simple & complex, respectively.

Conclusions- All models had good discrimination and acceptable calibration. Nevertheless, the highest calibration was observed for EuroSCORE model. This model was suggested for further assessment of isolated CABG surgery outcome at NMMC. Development of a more specific risk prediction model for NMMC was suggested for the future.

1. INTRODUCTION

Coronary artery bypass surgery (CABG), which has been used for more than 30 last years, remains the most accepted and established procedure for myocardial revascularization in multivessel coronary artery disease (1,2). Approximately three out of four patients are free from ischaemic events for five years following CABG resulting in improved quality of life (1). In general, after CABG procedure, about 96.5% of patients survive at least one month, 95% survive 1 year, 88% survive 5 years, 75% and 60% survive 10 years and 15 years respectively (1). The rate of hospital mortality of CABG surgery differs both across different centers and between the centers varying from 3% to 8.2 (1,3).

1.1. Risk stratification models

Quality assessment and its continuous improvement now is one of the requirements of good surgical practice (4). In cardiac surgery, as in many other types of surgery, operative or hospital mortality is used as an indicator of the quality of care. On the other hand, crude operative mortality is not always a good and valid measure of quality because it does not relate to the risk profile of patients receiving surgery (4). Without risk stratification, surgeons and hospitals treating high-risk patients will appear to have worse results than others which may influence their referral patterns, the allocation of resources, and even discourage treating high-risk patients (4). According to the EuroSCORE study group "risk stratification helps eliminate the bias against high-risk patients and informs patients and clinicians of the likely risk of death for a group of patients with a similar risk profile undergoing the proposed operation" (4).

In practice, the risk stratification results in development and implementation of different disease/condition and outcome specific risk models. Generally, these models analyze an individual array of risk factors to predict operative mortality or other outcome for a given patient (1). Individual patient risk factors are entered into the model, which then uses a statistical algorithm to calculate the probability that a given operative outcome will occur determining the *net impact* of all risk factors for an individual patient (5). The model does not predict "survival" or "death", but rather provides an estimate of the *probability* of operative death (5). Finally, any risk model should be both statistically valid and clinically interpretable as well as relevant to daily clinical decision making (6). The use of risk stratification models should form part of the basis on which the patient and surgeon decide whether to proceed or what is the most appropriate alternative method of treatment (1,4).

Currently there are several risk models used to predict both short- and long-term outcomes after CABG surgery. The Society of Thoracic Surgeons (STS) National Adult Cardiac Database is the largest voluntary clinical database in medicine with participants from about 522 sites from US and Canada (5,7). The database allows developing operative mortality and morbidity risk models for STS CABG procedures and updating them over time (8). This database was used to examine changes in the risk profile of patients undergoing isolated CABG and their outcomes during the decade from 1990 to 1999 (7). The results showed that, over time CABG patients were more likely to be older, of female gender, and have a history of smoking, diabetes mellitus, renal failure, hypertension, stroke, chronic ling disease, New York Heart Association functional class IV and three-vessel disease (7). Patients' predicted operative risk increased from 2.6% in 1990 to 3.4% in 1999 while the observed mortality decreased from 3.9% in 1990 to 3.0% in 1999 (7). These trends could be explained by improved coordination of care for CABG patient, formation of cardiac teams, technical

improvements in cardiac surgery, use of new pharmacological agents, and many other factors (7). Currently, STS uses a risk stratification model for CABG-only which includes 35 variables (1, 6). Among the other widely used models are European system for cardiac operative risk evaluation (EuroSCORE), Quality Measurement and Management Initiative (QMMI score), Predictive Score for Acquired Adult Heart Surgery (PARSONNET), UK Bayes model (simple and complex), ACC/AHA score etc (3,4, 9,10).

The usual procedure of model construction includes its development using one group of patients in terms of identification of significant risk factors (development dataset) and then its validation on another group of patients from the same patient population (validation dataset) (4-6). The development of these models required large samples of patients, comprehensive database systems, and great resources in terms of time, people, money, etc. That is why, in order to assess risk-stratified mortality rates, many centers prefer to use existing models. The first step in the latter process is the external validation of different models for the same patient population after which the model with the highest validity/predictive ability is chosen. Several validation studies showed that EuroSCORE was more acceptable for European population than PARSONNET, Ontario Province Risk Scores, and some other models developed based on US patient population (3, 11, 12) (Appendix 1). Among the other models demonstrated high predictive accuracy in European population was UK Bayes model (simple and complex) (12, 13) (Appendix 2).

1.2. Predictors of CABG operative mortality

The initial step to create any risk prediction model is to identify and measure the strength of different risk factors associated with the specific outcome. Different CABG mortality risk predictive models have both common risk factors and factors specific only to a given model. The selection of risk factors depends on the risk profile of the population, which was used to develop the model. Among the most consistent predictors of mortality after CABG surgery are advanced age, gender, priority of operation (elective versus urgent/emergent), prior heart surgery, left ventricular ejection fraction, recent myocardial infarction, number of major coronary arteries with significant stenoses, left main disease, as well as diabetes and renal disease (10,14-16). Women undergoing CABG have higher risk of postoperative mortality, than men, which was explained by their delayed referrals, high prevalence of known risk factors (diabetes mellitus, obesity, hypertension, etc), and smaller coronary vessels creating difficulties during the operation (1, 8,10). Cardiac factors like left ventricular ejection fraction, prior open-heart surgery, prior myocardial infarction, and valve diseases played a significant role on the outcome of CABG (4, 10). The difference in profiles of patients limited the use of the same risk prediction model across countries.

1.3. Quality Assurance Project at Nork Marash Medical Center

In March 2001 a Quality Assurance Project was launched at the Nork Marash Medical Center (NMMC) through a collaborative project between the Center and the American University of Armenia (AUA) Center for Health Services Research and Development (ANP project). NMMC provides a wide range of cardilogical and cardio surgical services to adult and pediatric populations. Quality improvement interventions were designed after the initial assessment of hospital functions which have resulted in positive changes in terms of hospital management and patient care practices (17). Since the establishment of NMMC in 1992, the number of cardiac operations had increased yearly. In 2003, more than 700 cardio surgical operations were performed, in which more than half were CABG procedure. No prior study

examined the 30-day operative mortality and the predictors of CABG procedure at NMMC. Such a data would create a basis for future assessment and improvement of quality of care. Such research could lead to the development of an appropriate risk model (highly predictive, typical for particular setting, disease, and patient population). At this stage it is essential and feasible to assess the profile of patient undergoing CABG surgery and validate an existing risk model to calculate risk adjusted 30-day operative mortality.

1.4. Study aims, research questions and objectives

The aim of the study was to assess the profile of patients who underwent isolated CABG procedure at NMMC in 2003 and validate an early mortality risk predictive model for this patient population.

The research questions addressed by the study were:

- What are the characteristics of patients who underwent isolated CABG procedure at NMMC in 2003?
- From three models (UK Bayes new simple, new complex, and EuroSCORE) what is the best one to predict the early operative (30-day) mortality risk of patients undergoing isolated CABG procedure at NMMC?

The objectives of the study were the following:

- Describe the profile of patient who underwent isolated CABG procedure at NMMC in 2003
- Evaluate the predictive accuracy of the UK Bayes new simple, new complex and EuroSCORE models for this patient population and recommend one for further use
- Measure the risk-stratified mortality of patients who underwent isolated CABG surgery at NMMC in 2003

2. METHODS

2.1. Study design

The study utilized retrospective review of NMMC surgical database, surgical medical histories, and first-visit structured encounter forms (SEF) of all cases undergone isolated CABG surgery from January 1, 2003 through December 31, 2003 at NMMC.

The ANP team and surgical team of NMMC discussed together the sources and some definitions of variables included in the study. The term ***isolated CABG*** does not include cases of myocardial revascularization performed in association with other procedures, such as cardiac valve replacement or repair, resection of left ventricular aneurysm or other cardiac operations. The term ***operative mortality*** was defined as death causally related to CABG surgery within the 30-day postoperative period. The cause of death was recorded from patient records or necroscopy examination when available.

2.2. Study protocol

All patients were examined by the cardiologists of Adult Cardiology Clinic (ACC) prior their admission to the Inpatient Clinic regardless of the type of visit (elective or urgent). After this encounter a structured encounter form (SEF) was completed for the patient. These SEFs,

surgical medical records, and the surgical database were reviewed for all the patients included in the sample. This was done to have accurate data and to handle missing data. Laboratory notes were used in cases when preoperative creatinine level was difficult to find in medical records.

Study instrument included all the variables needed for three models (Appendix 3).

2.3. Study population

Study population consisted from all patients who underwent isolated CABG surgery at NMMC from January 1, 2003 to December 31, 2003. Overall, data on 379 patients were collected and analyzed.

3. ETHICAL CONSIDERATIONS

The study involved secondary data analysis and possessed no risk for patients. Furthermore, the study was considered as a part of an internal evaluation process in the scope of Quality Assurance Project at NMMC. However, approval to access medical records, surgical database and SEFs was obtained prior to initiating the study from the hospital Medical Board. All records were reviewed inside the hospital to ensure patient confidentiality. In addition, the investigator undertook necessary actions to ensure confidentiality when there was a need to use patient names to retrieve information from all sources.

4. STUDY LIMITATIONS

One of the possible limitations is that the study was done retrospectively. This could result in some subjective judgment especially for the variables indicating the priority of operation in the UK Bayes models and the variable indicating critical preoperative state and emergency status for the EuroSCORE model. This limitation was handled using more than one source to retrieve patient information.

Pulmonary artery pressure was one of the variables of EuroSCORE model. It was excluded from the analyses because it was not directly measured at NMMC. EuroSCORE is an additive model and in this study was validated without this variable. Indeed, this could be a serious threat for study if the study population would consist of patients with cardiac valve disease, which was not the case.

5. DATA ANALYSES

SPSS 11.0 statistical software was used for single data entry. Range checking was done to ensure the accuracy of data. Data analyses were performed using SPSS 11.0, STATA 7.0 and Excel statistical software. As validation tool model discrimination and calibration were calculated. For model discrimination the area under the Receiver Operating Characteristic (ROC) curve was calculated and compared among the models. Model calibration was obtained by Hosmer-Lemeshow goodness-of-fit test with 8 degrees of freedom.

6. RESULTS

6.1. Administrative data

Overall 394 isolated CABG procedures were performed at NMMC from January 1, 2003 through December 31, 2003. The 30-day operative status was unknown for 9 patients who were from Georgia. All these patients were alive at the time of their discharge. In 7 cases the predicted mortality rates were not calculated because of missing data. All these records were excluded from the analyses. A total of 379 medical records were enrolled in the study.

6.2. Patient risk profile

Overall, 5 variables were needed for the UK Bayes simple model and 9 for the complex model, which were collected. From 18 EUROSCORE variables 17 were collected and entered into the model. The exception was the variable indicating pulmonary artery pressure. Some of the variables representing the risk profile of patient population were summarized in Table 1.

Table 1. Patient preoperative characteristics

Characteristics	Frequency (%)
Male	88.4
Ejection Fraction*	
Good (>=50%)	59.4
Fair (30-49 %)	39.3
Poor (< 30%)	1.3
Priority of operation*	
Elective	68.1
Urgent	31.6
Emergent	0.26
Previous heart operations	2.4
Left main disease	7.9
Diabetes	16.9
Acute or chronic renal failure	0.8
Hypertension	77.6
COPD**	2.1
Extracardiac arteriopathy	2.9
Neurological dysfunction	2.6

** Data was summarized as defined in the UK Bayes model (Appendix 2).*
***COPD - chronic obstructive pulmonary disease*

The mean age of patients was 55.2 years with a range from 28 to 80 (sd=9.5). Mean Body Surface Area was 1.93 kg/m^2 with a range from 1.09 to 2.37 (sd=0.17). Several patient characteristics were defined differently in UK and EuroSCORE models (ejection fraction, priority of operation, etc) and were presented as defined in one of the models. Most of the patients were males (88.4%), had good ejection fraction (68.1%) and underwent an elective procedure (68.1%). About one-third of patients had urgent operations. Prevalent preoperative comorbidities included hypertension (77.6%) and diabetes (16.9%). Few patients had COPD, extracardiac arteriopathy, neurological dysfunction, and acute or chronic renal failure.

6.3. Isolated CABG operative mortality

The observed crude operative mortality was 2.11% (8 out of 379 patients) with a confidence interval from 0.92% to 4.12%. The aggregate/mean predicted mortality rate by each model was calculated as the sum of estimated probabilities divided by the number of all patients. The mean overall predicted mortality was 1.53% (range 0.65 to 12.74%) by UK Bayes simple model, 1.46% (range 0.44% to 9.03%) by the complex model, and 2.36% (range 0.88% to 20.29%) by EuroSCORE model. There were no statistically significant differences between the observed mortality and predicted mortality rates from any of the models (Table 2).

The confidence intervals of SMRs were calculated using a normal approximation to the binomial distribution. There was no significant difference between the SMRs and 1, as evident from confidence intervals. This indicated that overall, all the three models gave accurate mortality estimates.

Table 2. Observed versus predicted operative death rate

Mortality predictive models	Mean predicted mortality rate, %	Observed vs predicted rates, *p* value	SMR (95% CI)
UK Bayes simple model	1.53	0.61	1.38 (0.6-2.73)
UK Bayes complex model	1.46	0.50	1.45 (0.63-2.86)
EuroSCORE model	2.36	0.82	0.90 (0.39-1.77)

** SMR – Standardized Mortality Ratio: the ratio of observed mortality to the aggregate predicted mortality, CI-Confidence Interval*

6.4. Validation of models: Discrimination and Calibration

Discrimination and calibration are the most frequently practiced methods to externally validate models. Discrimination is measured by ROC curve which "provides a measure of the model's ability to discriminate between those subjects who experience the outcome of interest versus those who do not" (18). It plots the probability of detecting the *true outcome (sensitivity)* and *false outcome (1-specificity)* for an entire range of possible cut points from 0 to 1 and calculates a C statistics corresponding to the total area under the ROC curve (18).

Table 3 presents the areas under the ROC curve for each model. The chi-square test comparing for significant differences between the areas under ROC curve of three models showed a non-significant result (*p* value was 0.44, Appendix 4). All models had excellent discrimination falling within the range from 0.8 to 0.9 (18).

Table 3. Discrimination of models

Predictive mortality models	C statistic (Area under the ROC curve)
UK Bayes simple model	0.869
UK Bayes complex model	0.864
EuroSCORE model	0.840

The next method of model diagnostic is the model calibration which assesses the concordance between deciles of observed and expected risk (18). The deciles of risk could be formed differently either by clinically relevant risk groups or equal size groups. For this study the latter method was applied: predicted risk was sorted from low to high and then was divided into ten equal size subgroups. To measure model calibration the Hosmer-Lemeshow

goodness-of-fit statistic was obtained. It yields a modified $\div^2$ statistic, where a small value is desired, yielding a non-significant *p* value difference between observed and estimated outcomes (18) (Appendix 4).

Table 4. Hosmer-Lemeshow goodness-of-fit test: ordered risk deciles

Predictive mortality models	Chi-square statistic	*p* value
UK Bayes simple model	13.39	0.10
UK Bayes complex model	8.60	0.38
EuroSCORE model	5.42	0.71

All models had acceptable calibration (non-significant *p* values) with the highest value for EuroSCORE model.

6.5. Risk-adjusted isolated CABG early mortality rate

Risk-adjusted mortality is an objective indicator of the cardiac surgery outcome. Ideally, to calculate the risk-adjusted mortality rate, the observed mortality rate is divided by the hospital's expected mortality rate. If the resulting ratio is larger than one, the hospital has a higher mortality rate than expected on the basis of the patient mix and visa versa if it is smaller than one. The ratio is then multiplied by the overall mortality rate obtained for the country/state in particular year to obtain the hospital's risk-adjusted rate. As NMMC is a unique cardio surgical center in Armenia, there is no nationwide mortality rate for isolated CABG to compare with. If one divides the observed mortality rate (2.11%) by the expected rate measured by EuroSCORE (2.36%) the result would be less than one, indicating good surgical outcome. Another way to present risk-adjusted mortality rate for isolated CABG at NMMC in 2003 is the use of clinically relevant risk strata. Table 5 summarized the results. EuroSCORE model was used to create clinically relevant risk strata (6).

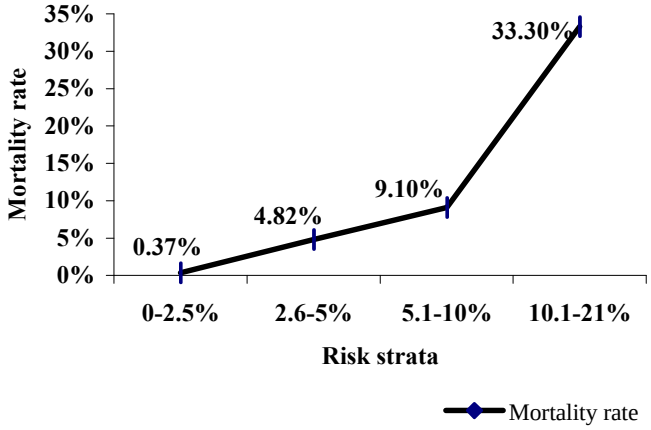
Table 5. Risk-adjusted mortality

Risk strata	Number of patients	Number who died	Mortality rate (%)
0-2.5%	271	1	0.369
2.6-5%	83	4	4.819
5.1-10%	22	2	9.090
10.1-21%*	3	1	33.300
Total	**379**	**8**	**2.11**

** There was no patient in the risk strata above 21%*

More than half of the patients were in the risk group from 0 to 2.5% which has the smallest mortality rate as well. The highest mortality rate 33.3% was observed in the highest risk strata. Figure 1 graphically displayed the increase in mortality rate with increasing patient predicted risk.

Figure 1. Risk-adjusted mortality



7. DISCUSSION

During the past decade when implemented the quality improvement programs have had a profound effect on improving mortality after CABG (7). Local analyses of processes and outcomes, and regional information exchange are tools that impact favorably on the outcomes following CABG. Validation of risk predictive models allows comparing the results from different hospitals and surgeons in a meaningful way which is an important goal for internal and external audit. The selection of an appropriate model should be done after detailed validation analyses of model.

The data was collected for all patients of NMMC who underwent isolated CABG procedure in 2003. The study involved 379 cases of isolated CABG. Separate analyses of patient preoperative risk profile showed that about two-thirds of patients suffered from hypertension, and 17% from diabetes. Only 2.1% of patients had COPD which was defined by EuroSCORE as long-term use of bronchodilators or steroids. Furthermore, the study population could have a casemix different from the reference populations of UK Bayes and EuroSCORE developmental models. Nevertheless, the study did not have the aim to compare the profiles of different patient populations and was directed only to validate the models.

Some of the variables were defined differently in comparing models. EuroSCORE model acknowledged 30-50% left ventricular EF as ‘moderate’ and less than 30% as ‘poor’. UK Bayes both models divided EF into the groups in the following way: good >=50%, fair=30-49% and poor<30%. Many patients who had 50% EF fell into different groups as defined by the models (‘moderate’ in EuroSCORE and ‘good’ in UK Bayes models). As at the NMMC EF 50% is accepted as good in patient preoperative profile this variable was summarized using the definition of UK Bayes model. Priority of operation was represented by one variable with three levels in UK Bayes model (elective, urgent, emergent), while EuroSCORE had separate variables of ‘emergency status’ and ‘critical preoperative sate’. In both models the selection of the extent of priority of operation could have some subjectivity because the data was collected retrospectively. For some patients, all three sources of information (SEFs, medical record, and surgical database) were carefully examined prior a final selection of priority of operation.

Assessing a model accuracy is a difficult task. Direct comparison of predicted and observed mortality of an individual patient is not a useful method to evaluate the predictive ability of a model. The area under the ROC curve is a more valuable summary statistical measure. An area of 1 suggests a perfect discrimination while a value of 0.5 shows no discrimination. The overall predictive abilities of the UK Bayes simple, complex and EuroSCORE models as indicated by the areas under ROC curves revealed excellent discrimination. All three values of C statistics were in the range from 0.8 to 0.9 (18). Moreover, the areas were not significantly different from each other.

Another measure of a model's predictive accuracy is its calibration. It proves the model's ability to predict outcome across different risk deciles. A model could have excellent discrimination but poor calibration overestimating or underestimating the risk in different deciles. For this study the risk of patients was sorted in ascending manner and divided into ten equal size groups. All the models have satisfactory calibration with *p* value of $\div^2$ statistic more than 0.5. Nevertheless, the highest value (p=0.71) was observed for EuroSCORE model. The high predictive ability of EuroSCORE model was proved in many other European countries as well (3,11).

Mean crude operative mortality of this patient population was 2.11% and was not significantly different from the models' predicted mean mortality rates. Standardized Mortality Ratios with the corresponding confidence intervals were calculated for all the models. Comparing with others, EuroSCORE model had the narrower confidence interval for SMR. Furthermore, this model was used to calculate risk-adjusted mortality risk of study population to reflect the quality of provided care. The results indicted that mortality rate was different across different clinically relevant risk strata. The observed death rate increased with the increase of patient predicted risk. The smallest mortality rate of 0.37% was observed in the lowest risk strata (0-2.5%) while the highest mortality rate was observed in the highest risk strata (10.1-21%) where out of 3 patients 1 died.

8. CONCLUSIONS/RECOMMENDATIONS

The goal of model validation is selection of a model with high predictive power. Among the investigated scores, the EuroSCORE yielded the highest predicted value in terms of model calibration and discrimination in this patient population. If properly used, it could predict early operative mortality of patients who undergo isolated CABG at NMMC. UK Bayes simple and complex models also shown excellent discrimination and acceptable calibration but could have more limited use. Individual patient risk calculation will enable the surgeon and hospital administration to make appropriate treatment decision and correct resource allocation.

Predictive models are powerful tools in surgical research. Accounting for the impact of all other important variables, the validated models could be used to determine the influence of a single factor on operative outcome. On the other hand, EuroSCORE or UK Bayes models could be used to internationally report coronary artery surgery results.

For continuous quality improvement, NMMC/ANP team should conduct more local self-assessment and self-improvement activities. Development of a more specific risk predictive model should be initiated because no risk adjustment model is better than the data upon which it is based. The derivation and validation data sets should be taken from NMMC

patients. The newly developed surgical database, which is more comprehensive than the previous one would make possible the construction of such a model. But this could be done only after four-five years highly depending on the number of surgical interventions. The new model would be then used mostly for internal audit while the validated EuroSCORE model for external audit.

REFERENCES

1. Gupta V, Grover V. Risk assessment and outcome after CABG. Annals of Cardiac Anaesthesia 2002;5:156-158.
2. Jones RH, Hannan EL, Hammermeister KE, et al. Identifcation of preoperative variables needed for risk adjustment of short-term mortality after coronary artery bypass graft surgery. J Am Coll Cardiol 1996;28:1478-1487.
3. Roques F, Nashef SAM, Michel P, at al. Does EuroSCORE work in Individual European countries? European Journal of Cardio-Thoracic Surgery 2000;18:27-30
4. Nashef SA, Roques F, Michael P, et al. Europian system for cardiac operative risk evaluation (EuroSCORE). Europian Journal of Cardio-thoracic Surgery 1999;16:9-13.
5. Edwards FH. Operative Risk Assessment In Cardiac Surgery. Jacksonville Medicine. October 2001.
6. Shroyer AL, Plomondon ME, Grover FL, Edwards FH. The 1996 coronary artery bypass risk model: The Society of Thoracic Surgeons Adult Cardiac National Database. Ann of Thorac Surg 1999;67:1205-8.
7. Ferguson TB, Hammill BG, Peterson ED, at al. A decade of change-risk profiles and outcomes for isolated coronary artery bypass grafting procedures, 1990-1999: A report from the STS National Database Committee and Duke Clinical Research Institute. Ann Thorac Surg 2002;73:480-90.
8. Shroyer AL, Coomba LP, Peterson ED, et al. The Society of Thoracic Surgeons: 30-day operative mortality and morbidity risk models. Ann Thorac Surg 2003;75:1856-65.
9. French Society of Anesthesia and Intensive Care. Scoring systems for ICU and surgical patients. Online. [Accessed 15 February, 2004]. Available from URL: http://www.sfar.org/s/article.php3?id_article=60
10. Coronary artery bypass graft surgery. A report of the American College of Cardiology/American Heart Association task force on practice guidelines. American College of Cardiology and American Heart Association, Inc. August 2001.
11. Geissler HJ, Hölzl P, Kuhn-Régnier F, at al. Risk stratification in hearth surgery: comparison of six score systems. Eur J Cardiothotac Surg 2000;17:400-406
12. Bridgwater B., Neve H., Hooper T., Jones M. Predicting operative risk for coronary artery surgery in the United Kingdom: a comparison of various risk prediction algorithms. Heart 1998;79:350-355
13. National adult cardiac surgical database report 2000-2001. Online [Accessed 4 March, 2004]. Available from URL: http://www.ctsnet.org/
14. Tu JV, Sykora K, Naylor CD. Assessing the outcomes of coronary artery bypass graft surgery: how many risk factors are enough? J Am Coll Cardiol 1997;30:317-1323.
15. Higgins TL, Estafanos FG, Loop FD, et al. ICU admission score for predicting morbidity and mortality risk after coronary artery bypass grafting. Ann Thorac Surg 1997;64:1050-1058
16. Ivanov J, Borger MA, David TE, at al. Predictive accuracy study: comparing a statistical model to clinicians' estimates of outcome after coronary bypass surgery. Ann Thorac Surg 2000;70:162-8.
17. Oksuzyan A, Demirchyan A, Thompson M. Evaluation of nurse-assessed work environment and job satsfaction at nork marash medical Center. Yerevan, Armenia: American University of Armenia, Center for Health Services Research and Development and Nork marash medical Center, 2003.
18. Hosmer DW, Lemeshow S. Applied logistics regression. Second edition. A Wiley-Interscience Publication. JohnWiley & SONS.INC. New York.2000

Appendix 1. European system for cardiac risk evaluation (EuroSCORE)

Factors	Criteria	Beta[1]	Definitions
Patient factors			
Age	Continuous	0.0666354	Age at the time of surgery
Gender	Female	0.3304052	Male, female
	Male		
COPD	Yes	0.4931341	Longterm use of bronchodilators or steroids for lung disease
	No		
Extracardiac arteriopathy	Yes	0.6558917	Anyone or more of the following: claudication, carotid occlusion or >50% stenosis, previous or planned intervention on the abdominal aorta, limb arteries or carotids
	No		
Neurological dysfunction	Yes	0.841626	Severely affecting ambulation or day-to-day functioning
	No		
Previous cardiac surgery	Yes	1.002625	Requiring opening of the pericardium
	No		
Serum creatinine> 200 µmol/ L	Yes	0.6521653	Serum creatinine >200m micromol/L preoperatively
	No		
Active endocarditis	Yes	1.101265	Patient still under antibiotic treatment for endocarditis at the time of surgery
	No		
Critical preoperative state	Yes	0.9058132	Any one or more of the following: ventricular tachycardia or fibrillation or aborted sudden death, preoperative cardiac massage, preoperative ventilation before arrival in the anaesthetic room, preoperative inotropic support, intraaortic balloon counterpulsation or preoperative acute renal failure (anuria or oliguria<10 ml/hour)
	No		
Cardiac factors			
Unstable angina	Yes	0.5677075	Rest angina requiring iv nitrates until arrival in the anaesthetic room
	No		
LVEF moderate	Yes	0.4191643	Left ventricular ejection fraction 30-50%
	No		

[1] Beta – regression coefficient of the factor. The Odds Ratio could be obtained by the exponentiation of beta.

Factors	Criteria	Beta[1]	Definitions
LVEF poor	Yes	0.5677075	Left ventricular ejection fraction < 30%
	No		
Recent myocardial infarction (< 90 days)	Yes	0.5460218	Myocardial infarction within 90 days before surgery
	No		
Pulmonary hypertension	Yes	0.7676924	Systolic PAP > 60 mmHg
	No		
Operation factors			
Emergency	Yes	0.7127953	Carried out on referral before the beginning of the next working day
	No		
Other than isolated C.A.B.G.	Yes	0.5420364	Major cardiac procedure other than or in addition to CABG
	No		
Surgery on thoracic aorta	Yes	1.159787	For disorder of ascending, arch or descending aorta
	No		
Postinfarct septal rupture	Yes	1.462009	Post myocardial infarction septal rupture
	No		

The European system for cardiac operative risk evaluation (EuroSCORE) was developed in 1998 to predict early mortality of European cardiac surgery patient population (4). It is a continuously up-to-date system. The model includes three groups of risk factors: patient-related, cardiac and operation-related factors. The model and its update are available from URL: ***http://www.euroscore.org/***

Appendix 2. UK Bayes Models (Simple & Complex)

Risk factor	Criteria	Odds Ratio	Definitions
Age	< 56 years old	0. 50	Age at the time of surgery; whole years rounded down.
	56-60 years old	0.53	
	61-65 years old	0.80	
	66-70 years old	1.09	
	71-75 years old	1.60	
	> 75 years old	2.75	
Body Surface Area	< 1.70 m^2	1.63	Calculated by the formula: (7.184^10-3) x (mass^0.425) x (height^0.725)
	0.70-1.89 m^2	1.16	
	1.90-2.39 m^2	0.86	
	> 2.39 m^2	0.66	
Ejection fraction	Good	0.64	Good>=50%, Fair=30-49% and Poor<30%
	Fair	1.18	
	Poor	3.88	
Priority	Elective	0.70	Elective: routine admission from the waiting list; procedure may be deferred without risk, Urgent: not scheduled for routine admission from the list but require surgery on the current admission for medical reasons; they cannot be sent home without surgery, (worsening, sudden chest pain, CHF, AMI, unstable angina or rest angina), Emergency: unscheduled with ongoing refractory cardiac symptoms; there should be no delay in surgical intervention irrespective of the time of day (pulmonary edema, carcinogenic shock), Salvage: requiring CPR en route to theatre or prior to anaesthetic; (Lump together Emergency and Salvage to give the "Emergency" risk group for Bayes)
	Urgent	1.22	
	Emergency	6.23	
Previous operations	None	0.93	Number of previous heart operations
	One or more	2.72	
Diabetes	No diabetes	0.95	Presence of oral / insulin / diet controlled diabetes
	Diabetes	1.29	
Hypertension	No hypertension	0.84	History of blood pressure > 140/90 mmHg on two occasions or lower if on medication
	Hypertension	1.18	
Left main stem disease	No LMS	0.86	Left main stem disease >50% diameter stenosis
	LMS	1.42	

Renal system	Dialysis	3.71	Dialysis = either Acute renal failure - dialysis (renal failure within 6 weeks of surgery necessitating any form of dialysis up to the time of surgery) or Chronic renal failure - dialysis (chronic renal failure on regular dialysis); Raised creatinine (Creatinine >200 micromol/L at the time of surgery)
	Raised creatinine	3.36	
	No renal disease	0.90	

UK Bayes models were created to predict in-hospital mortality in patients undergoing isolated CABG. It was developed 1999 and was tested in 1999 and 2000 separately (13). ***Simple model*** *includes 5 risk factors: age, body surface area, ejection fraction, priority of operation, and previous operations.* ***Complex model*** *includes all five factors of simple model plus four other factors: diabetes, hypertension, left main stem disease and renal system. The model is available from URL:* http://www.ctsnet.org/

Appendix 3. Study instrument

ID_ _ _

Name_____

1. Date of hospital admission _ _ / _ _ / _ _	**2. Date of discharge** _ _ / _ _ / _ _	**3. Date of surgery** _ _ / _ _ / _ _
4. Date of birth _ _ /_ _/ _ _	**5. Gender** 0. Male 1. Female	
6. Country of residency 1. Armenia 2. Georgia 3. Russia 4. Other	**7. Diagnosis at admission to ACC** 1. IHD 5. CHD 9. Other 2. HVD/rheumatic 6. Arrhythmia 3. HVD/non rheumatic 7. Cardiomyopathy 4. Hypertension 8. MI_____	
8. IHD 1. Stable angina **2. Unstable angina (CCS)**		**9. Cong. Heart Failure (NYHA)** 0. No 1. Yes
10. History of Miocardial infarction 0. No 1. Yes	**11. Cardiogenic shock** 0. No 1. Yes **11_1.** 1. Hemod.instability 2. Refractrory shock	**12. Arrhythmia** 0. No 1. Yes 1. Sust. VT/VF 3. AFib/Flutter 2. Heart block
13. Previous cardiac intervention (besides the last one) 0. No 1. Yes #____	**14. Previous cardiac surgery** 0. No 1. Yes #_____	**15. a. Weight**_____ **b.Height** _____ **16. BMI** _____ kg/m^2 **17. BSA** _____ m^2
18. Smoking status a. Present 0. No 1. Yes b. Past 0. No 1. Yes	**19. Physical activity** 1. Active 2. Moderate 3. Sedentary	**20. Drug allergy** 0. No 1. Yes **21. Other allergy** 0. No 1. Yes
22. Family history of IHD, SCD, or hypertension 0. No 1. Yes	**23. Diabetes mellitus** 0. No 1. Yes **23_1. Type of control** 1. Oral 2. Insulin 3. None 4. Diet	
24. Cerebrovascular accident 0. No 1. Yes	**25. Respiratory disease** 0. No 1. Yes	**26. Extracardiac arteriopathy** 0. No 1. Yes
27. Infectious endocarditis 0. No 1. Yes **27-1.** 1. Treated 2. Active	**28. Hypercholesterolemia (Chol. > 5.7mmol/L)** 0. No 1. Yes	**29. Hypertension** 0. No 1. Yes

30. Urogenital diseases 0. No 1. Yes	**31. Renal diseases** 0. No 1. Yes	**32. Preoperative creatinine (>200ì mol/L)** 0. No 1. Yes
33. Gastrointestinal diseases 0. No 1. Yes	**34. Other diseases** 0. No 1. Yes 34_a. _____	**35. Priority of operation** 1. Elective 2. Urgent 3. Emergent
36. Number of diseased coronary vessels 1. None 2. Two 3. One 4. Three	**37. Left main disease** (>50% stenosis) 0. No 1. Yes	**38. Ejection fraction** _____% 1. Good (>=50%) 2. Fair (30-49%) 3. Poor (<30%)
39. Postinfarct Septal Rupture 0. No 1. Yes	**40. CPB** 0. No 1. Yes	**41. IABP** 0. No 1. Yes
42. Length of ICU stay (hours) _____	**43. 30 day postoperative status** 1. Dead 0. Alive	**44. Postoperative (inhospital) complications** 0. No 1. Yes

Postoperative (inhospital) complications

45. Bleeding	No Yes	
46. Arrhythmia	No Yes	
47. Wound infection	No Yes	
48. Pneumonia	No Yes	
49. Stroke/TIA	No Yes	
50. MI	No Yes	
51. Reoperation	No Yes	
52. Other	No Yes	
52_a. Specify other_____		

53. Predicted death rate

a.UK Bayes Simple _____ % b.UK Bayes Complex _____ % c. EuroSCORE _____ %

Appendix 4. Data analyses: Receiver Operating Characteristic (ROC) curves

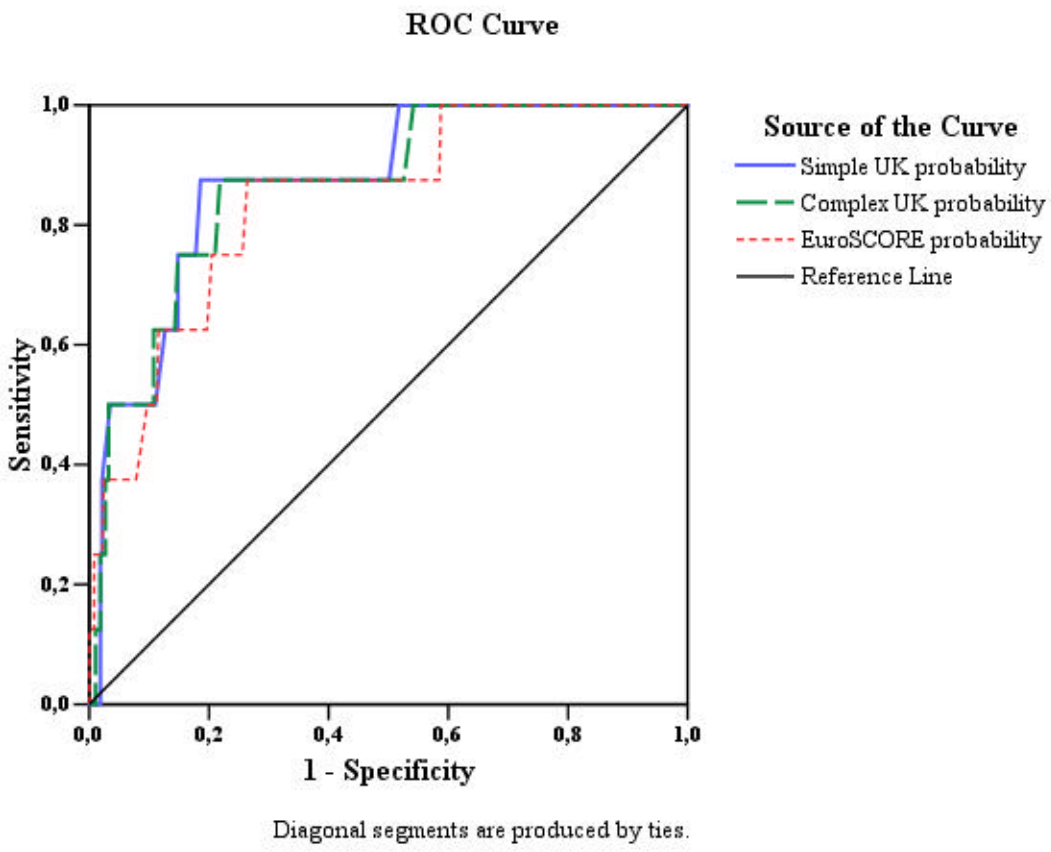
1. STATA output

. roccomp postop simple complex eur_scor, graph summary

	Obs	ROC Area	Std. Err.	-Asymptotic Normal-- [95% Conf. Interval]	
simple	379	0.8693	0.0588	0.75402	0.98453
complex	379	0.8637	0.0631	0.74003	0.98739
euroscore	379	0.8400	0.0699	0.70303	0.97689

Ho: area(simple) = area(complex) = area(euroscore)
chi2(2) = 1.64 Prob>chi2 = 0.4400

2. Graphical display of ROC curves



Appendix 5. Data analyses: Calibration of models

1. Formula for calculating Hosmer-Lemeshow goodness-of-fit statistic:

$$HL = \sum_{j=1}^{G} \frac{(o_j - n_j \bar{\pi}_j)^2}{n_j \bar{\pi}_j (1 - \bar{\pi}_j)}$$

where $\mathbf{o_j}$ is the number of positive observations in group j, $\mathbf{\delta_j}$ is the model's average predicted value in group j, and $\mathbf{n_j}$ is the size of the group. The Hosmer-Lemeshow statistic follows a chi-squared distribution with *G*- 2 degrees of freedom.

2. Summary table of Hostmer-Lemeshow goodness-of-fit test: ordered risk deciles*

Risk-deciles	**UK Bayes Simple model**		**UK Bayes Complex Model**		**EuroSCORE model**	
	Observed death	Aggregate expected mortality	Observed death	Aggregate expected mortality	Observed death	Aggregate expected mortality
1	0	0.65	0	0.52	0	0.88
2	0	0.7	0	0.61	0	1.2
3	0	0.88	0	0.76	0	1.33
4	0	1.03	0	0.91	0	1.34
5	1	1.1	1	0.99	1	1.56
6	0	1.3	0	1.1	0	1.99
7	0	1.5	0	1.4	0	2.3
8	0	1.8	1	1.7	2	3.4
9	3	2.3	2	2.3	1	3.8
10	4	3.98	4	4.17	4	6.4
HL statistic	**13.39**		**8.60**		**5.42**	
Prob (chi2)	**0.10**		**0.38**		**0.71**	

** After the predicted risk of death for each patient was arranged increasingly the population was divided into 10 groups of equal size.*