

Performance of EuroSCORE in CABG and off-pump coronary artery bypass grafting: single institution experience and meta-analysis

Alessandro Parolari^{1*}, Lorenzo L. Pesce², Matteo Trezzi¹, Claudia Loardi¹, Samer Kassem¹, Claudio Brambillasca³, Bruno Miguel⁴, Elena Tremoli¹, Paolo Biglioli¹, and Francesco Alamanni¹

¹Department of Cardiac Surgery, Unit for Clinical Research in Atherothrombosis, Centro Cardiologico Monzino IRCCS, University of Milan, Via Parea, 4, 20138 Milan, Italy; ²Department of Radiology, MC 2026, The University of Chicago, 5841 South Maryland Avenue, Chicago, IL, USA; ³Department of Anesthesia and Intensive Care, Centro Cardiologico Monzino IRCCS, Milan, Italy; and ⁴Department of Cardiovascular Surgery, Clermont-Ferrand University, Clermont-Ferrand, France

Received 22 September 2008; revised 24 November 2008; accepted 4 December 2008; online publish-ahead-of-print 13 January 2009

Aims

To assess EuroSCORE performance in predicting in-hospital mortality in on-pump coronary artery bypass grafting (CABG) and off-pump coronary artery bypass grafting (OPCAB).

Methods and results

Additive and logistic EuroSCORE were computed for consecutive patients undergoing CABG ($n = 3440$, 75%) or OPCAB ($n = 1140$, 25%) at our hospital from 1999 to September 2007. The areas under the receiver operating characteristic (ROC) curves (AUCs) were used to describe performance and accuracy. No difference in performance between CABG and OPCAB and between additive and logistic EuroSCORE (additive EuroSCORE AUCs of 0.808 and 0.779 for CABG and OPCAB, respectively; logistic EuroSCORE AUCs of 0.813 and of 0.773 for CABG and OPCAB, respectively) was found, although a marked tendency to overpredict mortality by both models was evident. A meta-analysis of previously published data was done, and a total of eight studies representing 19 212 and 5461 patients undergoing CABG and OPCAB, respectively, met inclusion criteria. Meta-analysis confirmed similar performance of EuroSCORE in CABG and OPCAB: estimated AUCs were 0.767 and 0.766 for CABG and OPCAB, respectively, with an estimated difference of 0.001 (95% CI -0.061 to 0.063).

Conclusion

Additive and logistic EuroSCORE algorithms performed similarly, and cumulative evidence suggests comparable performance in CABG and OPCAB procedures; both risk models, however, significantly overestimated mortality.

Keywords

EuroSCORE • CABG • OPCAB • ROC analysis • Meta-analysis

Introduction

In-hospital mortality is frequently used as an indicator of the quality of care in cardiac surgery, and EuroSCORE has gained wide popularity among risk-stratifying tools as it has been validated with good results in European, North American, and Asiatic Institutions.^{1–8}

Risk scoring systems, however, show their best performance when the pre-operative patient characteristics and treatment profiles are comparable with those on which the system was originated. For this reason, any risk scoring system can only be used reliably when its validity has been tested in the local patient

population and when surgical techniques and—more widely—patient management did not show substantial changes after the development of the risk score.⁹

Off-pump coronary artery bypass grafting (OPCAB) is now considered a surgical option to treat coronary artery disease; nowadays, ~20–25% of coronary bypass procedures are performed off-pump in the United States. The relatively recent introduction of OPCAB in clinical practice, which occurred sensibly after the development of currently used risk estimation models, raises the question whether these algorithms are appropriate for risk estimation in OPCAB.

*Corresponding author. Tel: +39 02 580021, Fax: +39 02 58011194, Email: alessandro.parolari@cardiologicomonzino.it

Published on behalf of the European Society of Cardiology. All rights reserved. © The Author 2009. For permissions please email: journals.permissions@oxfordjournals.org.

In this study, we have evaluated whether there are differences in EuroSCORE performance in estimating in-hospital mortality of patients undergoing standard coronary artery bypass surgery with the use of cardiopulmonary bypass on-pump coronary artery bypass grafting (CABG) and of patients undergoing OPCAB; this was done by first analysing our patients data and then by performing a systematic review of previously published data.

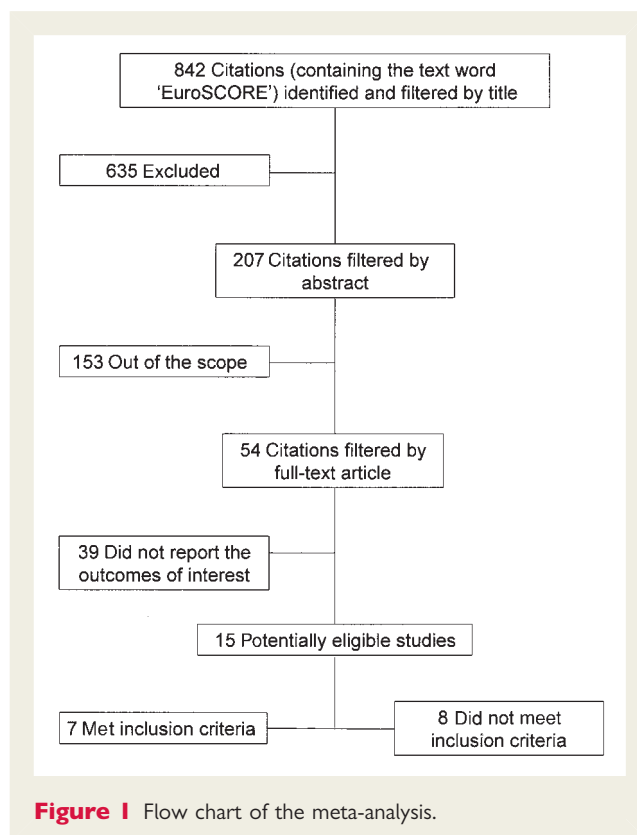
Methods

Single institution experience

This study includes 4580 consecutive patients who underwent isolated surgical myocardial revascularization at our hospital from January 1999 through September 2007; patients who had associated cardiac or non-cardiac procedures performed were excluded from the study. Pre-operative, operative, and in-hospital outcome data were prospectively collected in a computerized database—i.e. as part of standard clinical procedures, which is part of the reason why we had no missing data issues—whose use for research was approved by Centro Cardiologico Monzino Institutional Review Board because of the large number of patients and duration of the study. The requirement for individual informed consent was waived on the condition that the subjects' identities were hidden before analysing the data. Additive and logistic EuroSCORE were prospectively calculated in our database starting from 1999 and 2003, respectively, whereas logistic EuroSCORE was retrospectively computed for the earlier time periods (we have no reason to expect that this will introduce a significant bias). We assessed both additive and logistic EuroSCORE performances in risk prediction for patients undergoing CABG and OPCAB using the following approaches: (i) comparison of observed and predicted mortality figures with 95% confidence intervals in all patients, in patient undergoing CABG and OPCAB, and in patients stratified for risk quartiles;¹⁰ (ii) receiver operating characteristic (ROC) analysis^{11,12} using the model of DeLong *et al.*¹³ implemented on the University of Chicago ROC software (<http://xray.bsd.uchicago.edu/krl/>) to compare the areas under the ROC curves (AUCs) generated for additive and logistic EuroSCORE and for patients undergoing CABG and OPCAB, respectively; (iii) calibration analysis, assessed by the Hosmer–Lemeshow test.¹⁴ *P*-values <0.05 were considered significant.

Meta-analysis

To gain further insight into how EuroSCORE performs for OPCAB vs. CABG, a meta-analysis was done following the Meta-analysis Of Observational Studies in Epidemiology (MOOSE) guidelines.¹⁵ On 15 March 2008, two reviewers searched Medline (1950 to 1st week of March 2008), Embase (1966 to March 2008), PubMed (up to 15 March 2008), and The Cochrane Library including electronic links to related articles. The text string employed was the single word 'EuroSCORE'. In order to further reduce the probability of losing any major related study an electronic search of four major cardio-thoracic surgery journals in the electronic format (Interactive CardioVascular and Thoracic Surgery, The Annals of Thoracic Surgery, The European Journal of Cardiothoracic Surgery, The Journal of Thoracic and Cardiovascular Surgery, available at <http://ats.ctsnetjournals.org/search.dtl>) was performed; the journals were searched from January 1999 to March 2008 for the single word 'EuroSCORE' in the full text of all articles. The title of every paper was considered first, then selected abstracts were searched in order to identify reports about risk prediction by EuroSCORE for coronary bypass surgery. Full text of these papers was retrieved and searched for in-hospital or 30-day mortality



prediction by additive and/or logistic EuroSCORE in CABG and/or OPCAB. *Figure 1* reports the flow chart of paper selection.

Once papers were identified, a mandatory selection criterion for meta-analysis inclusion of each study was the presence of the assessment of discriminatory power of EuroSCORE by ROC analysis reporting the figures of merit and the dispersion parameters (standard error and/or 95% CI) necessary for the meta-analysis. Since the purpose of this meta-analysis is to gain insight into the general discriminatory characteristics of the EuroSCORE models for the two surgical modalities, we included only the area under the curve, instead of building a complete summary ROC curve. Moreover, an extremely large number of cases is necessary to make reliable statements about ROC curve shapes. Our analysis follows the work of McClish¹⁶ and Zhou.¹⁷ Finally, the QUADAS quality assessment tool for systematic reviews of diagnostic accuracy was applied to all selected studies.¹⁸

All calculations were done either with University of Chicago software, Excel (Microsoft Office Excel 2003, Microsoft Inc., Redmond, WA, USA) or using Mathematica (Mathematica 6, Wolfram Research Inc., Champaign, IL, USA). A more detailed description of statistical methods is reported in Appendix 1 (see Supplementary material online).

Results

Single institution experience

The clinical features of the CABG and OPCAB patient population are reported in *Table 1*. Of note, both additive and logistic average EuroSCORE values were slightly, although statistically significantly, higher in OPCAB patients. Overall in-hospital mortality was 0.83% (38/4580 patients, 95% CI 0.70–1.00%), and it was 0.84% (95% CI 0.70–1.04%) in CABG and 0.79% (95% CI 0.56–1.20%) in OPCAB,

Table 1 Prevalence of predictors in the study population

	CABG	OPCAB	P-value (CABG vs. OPCAB)
<i>n</i>	3440	1140	
Age			
Mean \pm SD	65 \pm 9	67 \pm 9	<0.001*
<60 years	28.1	22.2	
60–64 years	18.2	17.5	
65–69 years	23.2	19.2	<0.001*
70–74 years	18.5	19.2	
75+ years	12.0	21.9	
Female	18.2	18.2	0.976
Chronic pulmonary disease	8.1	11.0	0.003*
Extra-cardiac arteriopathy	4.6	9.4	<0.001*
Neurological dysfunction	1.3	2.6	0.002*
Previous cardiac surgery	0.9	7.3	<0.001*
Creatinine (mg/dL)			
Mean \pm SD	1.09 \pm 0.46	1.16 \pm 0.63	0.003*
Serum creatinine >200 mmol/L	0.9	7.3	<0.001*
Critical pre-operative state	2.1	1.9	0.736
Unstable angina	10.6	10.5	0.958
LVEF			
Mean \pm SD	57 \pm 10.7	57 \pm 10.6	0.830
LVEF 30–50%	21.6	21.3	
LVEF <30%	1.5	1.2	0.800
Recent myocardial infarct	19.9	18.9	0.477
Pulmonary hypertension	1.2	1.2	0.941
Emergency	3.2	1.8	0.020
Additive EuroSCORE	3.26 \pm 2.45	3.94 \pm 2.57	<0.001*
Logistic EuroSCORE	3.26 \pm 4.12	4.08 \pm 4.39	<0.001*
In-hospital mortality	0.84	0.79	0.863

*Statistically significant after the Holm–Sidak multiple testing correction.

respectively. Figure 2 reports the plots of observed vs. predicted mortality rates by risk quartiles in all, in CABG, and in OPCAB patients, and the ratio between observed vs. expected deaths. In addition, the predicted vs. observed mortality rates for additive and logistic EuroSCORE in the whole patient population, in CABG, and in OPCAB by risk quartiles are reported in Appendix 2, Table A1 (see Supplementary material online).

Data show a marked propensity to over-predict in-hospital mortality in all risk quartiles; this was true for the whole patient population, and also when patients were stratified by surgical strategy (CABG vs. OPCAB). The behaviour of both additive and logistic models was very similar in both CABG and OPCAB.

Table 2 shows the calibration and discrimination performance of additive and logistic EuroSCORE for different subcategories of patients. All models showed good calibration by the Hosmer–Lemeshow test, suggesting that they predict accurately on average. Also, the discriminatory power, assessed using ROC areas, was fair for almost all the models, with values around 0.8 for all subcategories, indicating that all the models could be useful for stratifying patients into risk groups for surgical management. Notably, only small differences in discrimination were

observed for both additive and logistic EuroSCORE for patients undergoing CABG and OPCAB, although with some degree of uncertainty as shown by relatively wide confidence intervals (Table 2).

Finally, no differences were detected in discrimination between additive and logistic EuroSCORE models in the whole patient population (delta AUC = –0.003, SE = 0.007, 95% CI –0.017 to 0.011, $P > 0.05$), in patients undergoing CABG (delta AUC = –0.006, SE = 0.009, 95% CI –0.023 to 0.011, $P > 0.05$), and in patients undergoing OPCAB (delta AUC = 0.006, SE = 0.015, 95% CI –0.023 to 0.0354, $P > 0.05$).

Meta-analysis of EuroSCORE performance in CABG and off-pump coronary artery bypass grafting

Due to the degree of uncertainty shown by the relatively wide confidence intervals displayed in the analysis just described, caused by the relatively low number of events in our experience, a meta-analysis of previously published data was performed. The search of the term EuroSCORE yielded 374, 315, and 409 hits in Pubmed, Medline, and EMBASE, respectively. The search of the

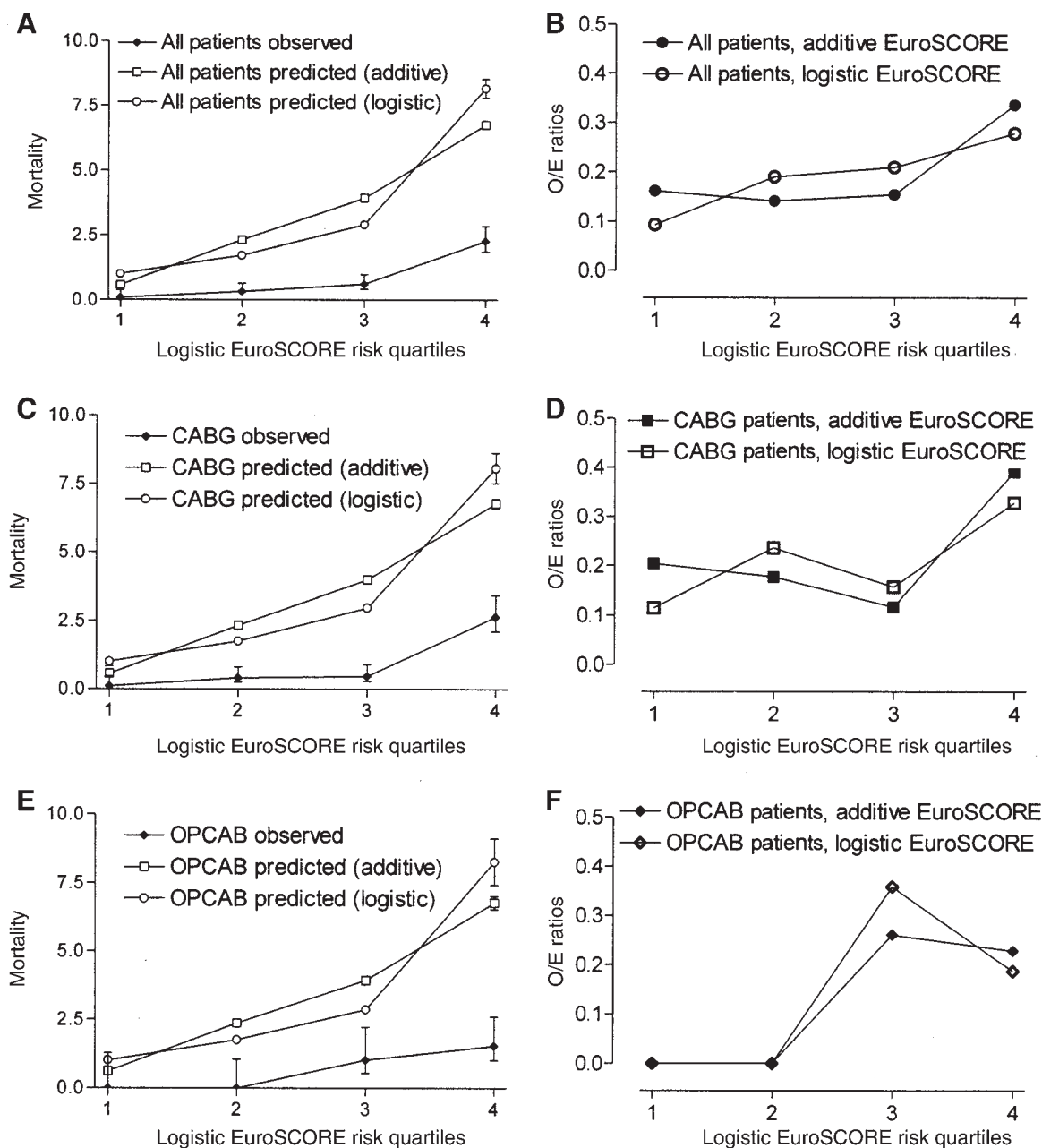


Figure 2 Plots of observed vs. predicted mortality rates by risk quartiles in all (A), in CABG (C), and in OPCAB (E) patients, and the ratio between observed vs. expected deaths (B, all patients; D, patients undergoing CABG; F, patients undergoing OPCAB).

term EuroSCORE in the full text of the four major cardiothoracic surgery journals yielded 573 hits. This led to a total of 842 papers available for analysis. The analysis of the abstracts and of the papers produced 15 studies,^{1,19–32} of whom seven fulfilled the selection criteria^{19–21,23,26,31,32} and eight did not;^{1,22,24,25,27–30} in addition, data from our study were included in the meta-analysis. The reasons for exclusion from the meta-analysis were: (i) the dispersion parameters (SE or CIs) were not reported;^{1,22,24,27,29} (ii) the type of surgical strategy (on- or off-pump) was not reported, or data were not analysed based on surgical strategy;^{22,24,25,29,30}

(iii) the dispersion parameters were very different from the values one would expect for a sample of that magnitude.²⁸ The features of the studies included in the meta-analyses are reported in Appendix 3, Tables B1 (additive EuroSCORE) and B2 (logistic EuroSCORE) (see Supplementary material online).

Table 3 reports the meta-analysis results comparing EuroSCORE AUCs in CABG and OPCAB; the AUC and SE from the logistic EuroSCORE was used when available, otherwise the additive one was employed. This decision takes into account the evidence that the two methods have nearly identical discriminatory power (all the

Table 2 Calibration and discrimination of additive and logistic EuroSCORE for different subcategories of patients

	ROC			Calibration		Unpaired ROC comparison
	AUC	95% CI	SE	P-value (χ^2 test, Hosmer–Lemeshow)		P-value
Additive EuroSCORE						
All patients (n = 4580)	0.802	0.730–0.873	0.037	0.96		0.78
CABG patients (n = 3440)	0.808	0.723–0.892	0.043	0.94		
OPCAB patients (n = 1140)	0.779	0.643–0.915	0.073	0.80		
Logistic EuroSCORE						
All patients (n = 4580)	0.805	0.735–0.875	0.036	0.25		0.75
CABG patients (n = 3440)	0.813	0.732–0.895	0.042	0.41		
OPCAB patients (n = 1140)	0.773	0.634–0.911	0.075	0.51		

Table 3 Meta-analysis of studies reporting discriminatory performance of additive and logistic EuroSCORE in CABG and OPCAB

	Patients	Events	Mortality % (SE)	AUC	95% CIs (SE)
CABG					
Antunes <i>et al.</i> ¹⁹	4567	44	0.96 (0.33)	0.754	0.68–0.83 (0.04)
Asimakopoulos <i>et al.</i> ²¹	4654	152	3.27 (0.26)	0.76	0.72–0.80 (0.02)
Biancari <i>et al.</i> ²³	1098	5	0.46 (0.26)	0.856	0.67–1.00 (0.09)
Farrokhyar <i>et al.</i> ²⁶	1693	26	1.54 (0.41)	0.81	0.71–0.90 (0.05)
Parolari (current study)	3440	29	0.84 (0.42)	0.813	0.73–0.90 (0.04)
Toumpoulis <i>et al.</i> ³¹	3760	103	2.74 (0.47)	0.75	0.71–0.80 (0.02)
Total	19 212	359	1.87 (0.098)	0.767	0.742–0.792 (0.13)
OPCAB					
Al-Ruzzeh <i>et al.</i> ²⁰	1907	26	1.36 (0.27)	0.75	0.64–0.85 (0.05)
Farrokhyar <i>et al.</i> ²⁶	1657	30	1.81 (0.33)	0.79	0.71–0.88 (0.04)
Parolari (current study)	1140	9	0.79 (0.26)	0.773	0.63–0.91 (0.07)
Youn <i>et al.</i> ³²	757	10	1.32 (0.41)	0.71	0.55–0.87 (0.08)
Total	5461	75	1.37 (0.16)	0.766	0.709–0.823 (0.029)

Estimate for delta AUC (CABG–OPCAB)=0.001 (95% CI –0.061 to 0.063, SE 0.032).

studies that attempted a comparison found truly negligible differences). It is noteworthy mentioning that this assumption does not imply that the two methods have the same predictive power or clinical value, but it simply acknowledges the fact that the AUC values do not differ. In our opinion, it would represent a bigger bias to ignore a study because it used the additive EuroSCORE model.

The samples showed no evidence of heterogeneity, which suggests that applying a meta-analysis approach to the data can provide reproducible estimates—see online supplementary Appendix 1 for details. ROC meta-analysis estimates for AUC are 0.767 (SE = 0.013) for CABG and 0.766 (SE = 0.029) for OPCAB, respectively, with an estimated difference between CABG and OPCAB AUCs of 0.001 (95% CI –0.061 to 0.063). Therefore, the difference in EuroSCORE discriminatory power is expected to be at most around 0.06. Finally, results from the application of the QUADAS 14-question quality assessment tool revealed that all the studies scored positively on at least 50% of the questions.

Discussion

Risk prediction and stratification models play an important role in current cardiac surgical practice, as health authorities, hospitals, and individuals (such as medical practitioners and patients) are paying more attention to objective risk-adjusted predictions of mortality after cardiac surgery. These models allow more meaningful comparisons of outcomes among institutions and surgeons by adjusting for different case-mix; they are also useful in surgical decision-making, pre-operative patient education and consent, and quality assurance measures. They can detect and quantify differences and changes in the risk profiles of patients undergoing cardiac surgery. Furthermore, risk prediction allows more objective balancing of potential risks and benefits for individual patients.³³

Risk stratification algorithms have reduced applicability when used in different patient populations from the ones they were formulated on. Models developed in the USA, for instance, may not

predict satisfactorily clinical outcome in European populations. Moreover, some of the currently used risk models do not reflect improved surgical techniques and post-operative patient management advances which occurred in recent times.

Additive EuroSCORE was first introduced in 1999,³ but data were collected from September 1995 to December 1995 in 128 European centres for a total of 19 030 patients available for final analysis;³⁴ in 2003, based on the same database, logistic EuroSCORE was introduced to improve the predictive performance for high-risk and very high-risk patients;³⁵ this new risk stratification tool uses a more complex algorithm to derive risk from the same pre-operative and operative risk factors.

Coronary bypass surgery without the use of cardiopulmonary bypass has been performed sporadically since the beginning of the bypass surgery era in the late sixties, but the use of this strategy gained considerable popularity between 1997 and 2000, then levelled off, e.g. accounting for 28% of all isolated CABG cases in New York between 2001 and 2004. Therefore, a consistent clinical use of OPCAB occurred after EuroSCORE was developed and no validation of the predictive performance of EuroSCORE between CABG and OPCAB was done. Also, of interest is the question whether additive and logistic models could have similar performances in CABG and, especially, OPCAB patients.

To our knowledge, few studies have addressed the performance of EuroSCORE in predicting outcome after off-pump coronary bypass surgery. Al-Ruzzeh et al.²⁰ in 2003 reported the results of a multicentre study examining the performance of additive EuroSCORE on 1907 OPCAB procedures with 1.3% in-hospital mortality, showing that the discriminatory power of additive EuroSCORE was 0.75 (95% CI 0.64–0.85); in this paper, no CABG group was available as comparison group, and for this reason no comparison between the two surgical strategies could be done. Berman et al.¹ report similar AUCs for patients undergoing CABG and OPCAB (ROCs of 0.76 and 0.74 in CABG and OPCAB, respectively) concerning additive EuroSCORE, but no formal comparison was made between these two groups. More recently, Youn et al.³² reported fair discrimination (AUC of 0.72 and 0.71 for additive and logistic EuroSCORE, respectively) for both additive and logistic EuroSCORE in 757 consecutive patients who underwent OPCAB between 2002 and 2006; also in this case, no CABG data were available as a control group. Finally, Farrokhlyar et al.²⁶ reported good discrimination for logistic EuroSCORE in risk prediction for patients submitted either to CABG (AUC of 0.81, 95% CI 0.71–0.90) or to OPCAB (AUC of 0.79, 95% CI 0.71–0.88), but no formal comparison between CABG and OPCAB discriminatory power was done, nor the performance of the additive model was considered.

Our study adds to the knowledge to date reasonable evidence that both additive and logistic EuroSCORE have similar performances, on average, in CABG and OPCAB; they are able to correctly stratify the risk for patients undergoing coronary bypass surgery, performed both on- and off-pump, but both risk scores sensibly overestimate the risk in CABG and in OPCAB, even when patients are stratified for risk. This occurs irrespective of the surgical strategy adopted and the discrepancy between observed and predicted mortality rates is particularly high in patients at low-medium risk. For this reason, caution should be exercised when comparing

in-hospital mortality directly against both additive and logistic EuroSCORE or in case of use of the models to compare results between hospitals or surgeons with different case-mix.

The observation that both additive and logistic average EuroSCORE values were slightly higher in OPCAB compared with CABG patients may be due to an almost double incidence of emergency operations in CABG patients. Indeed, this is not surprising given the different nature of the two techniques. It should be noted, however, that the purpose of this meta-analysis is to determine whether the predictive power of the models is similar for the two surgical approaches in the specific populations where they are utilized and not whether the two patient populations are identical. Generally speaking, we are addressing the question whether or not surgeons who are accustomed to EuroSCORE with CABG can confidently use the same score to OPCAB patients. In fact, given the different nature of the two techniques, it is to be expected that the patient-populations would be at least slightly different, as observed here.

Data from this study confirm and extend evidence from recent studies that have questioned the accuracy of both additive and logistic EuroSCORE algorithms in early mortality prediction reporting overprediction for both models,^{22,30,36} even when only patients undergoing standard on-pump coronary bypass surgery are considered.

The reasons for the results shown in this study are unclear and probably multi-factorial. The actual epidemiology of ischaemic heart disease and co-morbid conditions may be different from that of the population on which EuroSCORE was derived and computed. This could be due to real differences in the prevalence of disease states. It is also possible, and this is in our opinion the most likely explanation, that this is an artefact due to the impact of probably over time improved access to health care and to the management of these conditions in a heavily government-subsidised healthcare system, and to an early and aggressive investigation and management of these diseases. Furthermore, as the EuroSCORE was developed on data from patients in 1995, EuroSCORE models may not reflect anymore current cardiac surgical practice. In addition, patients undergoing coronary bypass surgery are increasingly becoming older and somehow fitter due to decreased prevalence of smoking, increased attention placed on lifestyle, and more aggressive medical treatment (e.g. statins) of cardiovascular diseases; this implies that the relative weight of increasing age and other risk factors in the EuroSCORE models may be actually sensibly different with respect to the estimates done for EuroSCORE models computation.

Limitations of the study

In this study, we used in-hospital mortality—defined as death occurring anytime during hospitalization for surgery—as indicator of early mortality. Another important indicator of performance in coronary bypass surgery is 30-day mortality. Although each of these indicators has advantages and disadvantages for precise risk assessment, previous studies have shown very similar performances of these two indicators, e.g. both similarly underestimate the early risk of death.³⁷

The estimation of EuroSCORE performance in particular subsets of patients (e.g. high-risk or low ejection fraction) would

have been of great interest to us; unfortunately, this analysis could not be performed because of the limited information existing and the relatively low numbers of patients available that precluded any chance of credible and reproducible subanalyses.

Conclusions

Our study has shown that EuroSCORE performs similarly in CABG and OPCAB patients and that the performances of the additive and logistic models are not different in these settings; in addition, the AUCs obtained in the ROC analysis suggest a fair discrimination of both models for CABG and OPCAB.

On the other hand, both models overpredict mortality at all risk levels; these findings should give cause for caution when using additive or logistic EuroSCORE as a risk-adjustment model for comparing outcomes between hospitals or surgeons, as it would be easy to gain false reassurance by comparing the observed mortality with that predicted by the algorithms. This suggests the need of EuroSCORE recalibration or even re-engineering in order to provide a more accurate and less variable scoring system for risk prediction.

Supplementary material

Supplementary material is available at *European Heart Journal* online.

Funding

L.L.P. gratefully acknowledges support from NIH contract HHSN267200700039C (Charles E. Metz, PI) for the development of part of the software used in the analysis (available as free download or by request at <http://xray.bsd.uchicago.edu/krll/>).

Conflict of interest: none declared.

References

- Berman M, Stamler A, Sahar G, Georgiou GP, Sharoni E, Brauner R, Medalion B, Vidne BA, Kogan A. Validation of the 2000 Bernstein-Parsonnet score versus the EuroSCORE as a prognostic tool in cardiac surgery. *Ann Thorac Surg* 2006;**81**: 537–540.
- Geissler HJ, Holzl P, Marohl S, Kuhn-Regnier F, Mehlhorn U, Sudkamp M, Rainer de Vivie E. Risk stratification in heart surgery: a comparison of six scoring systems. *Eur J Cardiothorac Surg* 2000;**17**:400–406.
- Nashef SA, Roques F, Michel P, Gauducheau E, Lemeshow S, Salamon R, the EuroSCORE study group. European system for cardiac operative risk evaluation (EuroSCORE). *Eur J Cardiothorac Surg* 1999;**16**:9–13.
- Kawachi Y, Nakashima A, Tushima Y, Arinaga K, Kawano H. Risk stratification analysis of operative mortality in heart and thoracic aorta surgery: comparison between Parsonnet and EuroSCORE additive model. *Eur J Cardiothorac Surg* 2001;**20**:961–966.
- Nilsson J, Algotsson L, Hoglund P, Luhrs C, Brandt J. Early mortality in coronary bypass surgery: the EuroSCORE versus The Society of Thoracic Surgeons risk algorithm. *Ann Thorac Surg* 2004;**77**:1235–1239.
- Nashef SA, Roques F, Hammill BG, Peterson ED, Michel P, Grover FL, Wyse RK, Ferguson TB. Validation of European System for Cardiac Operative Risk Evaluation (EuroSCORE) in North American cardiac surgery. *Eur J Cardiothorac Surg* 2002;**22**:101–105.
- Gogbashian A, Sedrakyan A, Treasure T. EuroSCORE: a systematic review of international performance. *Eur J Cardiothorac Surg* 2004;**25**:695–700.
- Nilsson J, Algotsson L, Hoglund P, Luhrs C, Brandt J. Comparison of 19 pre-operative risk stratification models in open-heart surgery. *Eur Heart J* 2006;**27**: 867–874.
- Ivanov J, Tu JV, Naylor CD. Ready-made, recalibrated, or remodeled? Issues in the use of risk indexes for assessing mortality after coronary artery bypass graft surgery. *Circulation* 1999;**99**:2098–2104.
- Omar RZ, Ambler G, Royston P, Eliahoo J, Taylor KM. Cardiac surgery risk modeling for mortality: a review of current practice and suggestions for improvement. *Ann Thorac Surg* 2004;**77**:2232–2237.
- Wagner RF, Metz CE, Campbell G. Assessment of medical imaging systems and computer aids: A tutorial review. *Acad Radiol* 2007;**14**:723–748.
- Pepe MS. *The Statistical Evaluation of Medical Tests for Classification and Prediction*. Oxford, New York: Oxford University Press; 2004.
- DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;**44**:837–845.
- Lemeshow S, Hosmer DW. A review of goodness of fit statistics for use in the development of logistic regression models. *Am J Epidemiol* 1982;**115**:92–106.
- Stroup DF, Berlin JA, Morton SC, Olkin I, Williamson GD, Rennie D, Moher D, Becker BJ, Sipe TA, Thacker SB, the MOOSE Group. Meta-analysis of observational studies in epidemiology: a proposal for reporting. Meta-analysis Of Observational Studies in Epidemiology (MOOSE) group. *JAMA* 2000;**283**:2008–2012.
- McClish DK. Combining and comparing area estimates across studies or strata. *Med Decis Making* 1992;**12**:274–279.
- Zhou XH. Empirical Bayes combination of estimated areas under ROC curves using estimating equations. *Med Decis Making* 1996;**16**:24–28.
- Whiting P, Rutjes A, Reitsma J, Bossuyt P, Kleijnen J. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Med Res Methodol* 2003;**3**:25.
- Antunes PE, Eugenio L, Ferrao de Oliveira J, Antunes MJ. Mortality risk prediction in coronary surgery: a locally developed model outperforms external risk models. *Interact Cardiovasc Thorac Surg* 2007;**6**:437–441.
- Al-Ruzzeh S, Asimakopoulos G, Ambler G, Omar R, Hasan R, Fabri B, El-Gamel A, DeSouza A, Zamvar V, Griffin S, Keenan D, Trivedi U, Pullan M, Cale A, Cowen M, Taylor K, Amrani M. Validation of four different risk stratification systems in patients undergoing off-pump coronary artery bypass surgery: a UK multicentre analysis of 2223 patients. *Heart* 2003;**89**:432–435.
- Asimakopoulos G, Al-Ruzzeh S, Ambler G, Omar RZ, Punjabi P, Amrani M, Taylor KM. An evaluation of existing risk stratification models as a tool for comparison of surgical performances for coronary artery bypass grafting between institutions. *Eur J Cardiothorac Surg* 2003;**23**:935–941.
- Bhatti F, Grayson AD, Grotte G, Fabri BM, Au J, Jones M, Bridgewater B, North West Quality Improvement Programme in Cardiac Interventions. The logistic EuroSCORE in cardiac surgery: how well does it predict operative risk? *Heart* 2006;**92**:1817–1820.
- Biancari F, Kangasniemi OP, Luukkonen J, Vuorisalo S, Satta J, Pokela R, Juvonen T. EuroSCORE predicts immediate and late outcome after coronary artery bypass surgery. *Ann Thorac Surg* 2006;**82**:57–61.
- Chen CC, Wang CC, Hsieh SR, Tsai HW, Wei HJ, Chang Y. Application of European system for cardiac operative risk evaluation (EuroSCORE) in coronary artery bypass surgery for Taiwanese. *Interact Cardiovasc Thorac Surg* 2004;**3**: 562–565.
- D'Errigo P, Seccareccia F, Rosato S, Manno V, Badoni G, Fusco D, Perucci CA, the Research Group of the Italian CABG Outcome Project. Comparison between an empirically derived model and the EuroSCORE system in the evaluation of hospital performance: the example of the Italian CABG Outcome Project. *Eur J Cardiothorac Surg* 2008;**33**:325–333.
- Farrokhhyar F, Wang X, Kent R, Lamy A. Early mortality from off-pump and on-pump coronary bypass surgery in Canada: a comparison of the STS and the EuroSCORE risk prediction algorithms. *Can J Cardiol* 2007;**23**:879–883.
- Karabulut H, Toraman F, Alhan C, Camur G, Evrenkaya S, Dagdelen S, Tarcan S. EuroSCORE overestimates the cardiac operative risk. *Cardiovasc Surg* 2003;**11**: 295–298.
- Sergeant P, de Worm E, Meyns B, Wouters P. The challenge of departmental quality control in the reengineering towards off-pump coronary artery bypass grafting. *Eur J Cardiothorac Surg* 2001;**20**:538–543.
- Yap CH, Mohajeri M, Ihle BU, Wilson AC, Goyal S, Yii M. Validation of Euroscore model in an Australian patient population. *ANZ J Surg* 2005;**75**:508–512.
- Yap CH, Reid C, Yii M, Rowland MA, Mohajeri M, Skillington PD, Seevanayagam S, Smith JA. Validation of the EuroSCORE model in Australia. *Eur J Cardiothorac Surg* 2006;**29**:441–446.
- Toumpoulis IK, Anagnostopoulos CE, DeRose JJ, Swistel DG. Does EuroSCORE predict length of stay and specific postoperative complications after coronary artery bypass grafting? *Int J Cardiol* 2005;**105**:19–25.
- Youn YN, Kwak YL, Yoo KJ. Can the EuroSCORE predict the early and mid-term mortality after off-pump coronary artery bypass grafting? *Ann Thorac Surg* 2007;**83**: 2111–2117.
- Kolh P. Importance of risk stratification models in cardiac surgery. *Eur Heart J* 2006;**27**:768–769.
- Roques F, Nashef SAM, Michel P, Gauducheau E, de Vincentis C, Baudet E, Cortina J, David M, Faichney A, Gabrielle F, Gams E, Harju la A, Jones MT,

- Pinna Pintor P, Salamon R, Thulin L. Risk factors and outcome in European cardiac surgery: analysis of the EuroSCORE multinational database of 19030 patients. *Eur J Cardiothorac Surg* 1999;15:816–823.
35. Michel P, Roques F, Nashef SA, the EuroSCORE Project Group. Logistic or additive EuroSCORE for high-risk patients? *Eur J Cardiothorac Surg* 2003;23:684–687.

36. Jin R, Grunkemeier GL, Providence Health System Cardiovascular Study Group. Does the logistic EuroSCORE offer an advantage over the additive model? *Interact Cardiovasc Thorac Surg* 2006;5:15–17.
37. Osswald BR, Blackstone EH, Tochtermann U, Thomas G, Vahl CF, Hagl S. The meaning of early mortality after CABG. *Eur J Cardiothorac Surg* 1999;15:401–407.

CARDIOVASCULAR FLASHLIGHT

doi:10.1093/eurheartj/ehn411

Online publish-ahead-of-print 10 September 2008

Progression of mitral annulus calcification to caseous necrosis of the mitral valve: complementary role of multi-modality imaging

Ron Blankstein^{1,2*}, Ronen Durst², Michael H. Picard², and Ricardo C. Cury¹

¹Cardiac MR PET CT Program, Department of Radiology, Massachusetts General Hospital, 165 Cambridge Street, Suite 400, Boston, MA 02114, USA and

²Division of Cardiology (Medicine), Massachusetts General Hospital, 55 Fruit St., Boston, MA 02114, USA

* Corresponding author: Tel: +1 3124049949, Fax: +1 617 724-5351, Email: rblankstein@partners.org

A 66-year-old male was found on echocardiography to have an echogenic discrete mass in the region of posterior mitral valve annulus (Panels A–D). In order to evaluate this area further, a cardiac MRI (Panels E and F) was performed and demonstrated a 20 × 20 mm mass in the posterior left ventricular myocardium between the coronary sinus and posterior mitral valve leaflet. No evidence of vascular flow was seen within the mass, and its tissue characteristics—notably low signal on proton density-weighted fast-spin echo—were most consistent with calcium, thus confirming the diagnosis of prolific mitral annular calcification.

Around the same time, a non-gated contrast-enhanced CT (Panel G) demonstrated a homogeneous hyperdense mass in the area of the inferior mitral valve annulus with 300 Hounsfield unit (HU). Five years later, a repeat CT for cancer surveillance (Panel H) showed the transformation of this area into a more heterogeneous calcification pattern notable for decreased signal within the centre of the mass (70 HU). This pattern is consistent with caseous necrosis, a rare calcification pattern which has central necrosis, and on pathology appears pale or caseous (i.e. ‘cheese like’).

This case illustrates the complementary role of different imaging modalities. Whereas echocardiography was used to initially detect this mass and to assess its functional significance, the use of cardiac MRI was used to exclude other potential aetiologies and to confirm the presence of calcium. Conventional CT imaging further confirmed the calcified nature of the mass and most optimally demonstrated the transformation into caseous necrosis.

