Validation of the Euroscore on Cardiac Surgery Patients in Nairobi

Awori Mark, Mehta Nikita, Mitema Fred, Mwangi Jimmy, Mjahid Hassan, Oloo Paul
School of Medicine, University of Nairobi

Correspondence to: Dr. Mark Awori. P.O Box 14677-00800 Nairobi. Email: mnawori@yahoo.com

Abstract

Background: The Additive Euroscore (AE) predicts outcomes in cardiac surgical procedures performed on cardiopulmonary bypass. It's been widely used in developed nations but it's applicability in Kenya is unknown. Our objective was to determine its applicability at Kenyatta National Hospital (Kenya). **Methods:** A retrospective study was carried out between 1st January 2011 and 31st December 2015. Risk factor prevalence was compared with that of the AE derivation population. The AE was calculated; discrimination was determined by receiver operator curve analysis. **Results:** Of 109 patients, significant differences (Kenyan vs. AE derivation) were found in the prevalence of pulmonary y hypertension (58.7% vs. 2%) and isolated coronary artery bypass graft surgery (4.6% vs. 65%). Only double valve replacement was a risk factor for operative mortality; odds ratio 5.98 (1.83 to 19.49). The area under curve (AUC) for the AE was 0.59. **Conclusion:** The AUC for the AE implies poor discrimination in our population. Significant differences in the risk factor profile between our study population and the AE score derivation population may have contributed to this. Our findings suggest that the AE may not be applicable to patients in Kenya. We recommend that a local risk scoring system be developed.

Key words: Euroscore, Validation, Kenya

Introduction

The aim of surgery is to improve patient health while minimizing patient exposure to procedure related morbidity and mortality; 'evidence based medical practice' (EBMP) is employed to facilitate this (1-3). Crude operative mortality does not take into account the pre-operative mortality risk factor profile of patients; EBMP dictates that 'risk adjusted operative morality' be employed to assess surgical outcome (3, 4). The use of risk stratification for the prediction of surgical outcomes is commonly practiced in most surgical disciplines. Risk stratification systems provide information that influences the choice of intervention, facilitates the process of obtaining informed consent and delivers a mechanism to measure surgical performance (5). Preoperative risk factors have been shown to affect surgical outcome in cardiac surgery (6). Risk stratification systems that incorporate only preoperative factors are particularly useful as they facilitate decision making before any intervention has been administered. For instance, by predicting the risk of operative mortality, patient and surgeon can opt for a procedure with a mutually acceptable mortality. Best practice requires that the highest possible surgical outcomes be obtained. In this regard, predicted mortalities using risk stratification systems can be used to determine if the observed mortality lies within the parameters of accepted quality. Several preoperative factors have been shown to affect surgical outcome in cardiac surgery and preoperative surgical risk prediction scores incorporating these factors have been developed (6). Many such scores were developed on specific patient populations and their applicability may not be

generalizable (7). As a result, cardiac surgical services have assessed the external validity of such scores and where necessary have recalibrated original scores or developed their own local scores to better predict outcomes in their patient population (8, 9). The additive Euroscore (AE) has been widely used in developed nations; of particular note, a low preoperative cardiac ejection fraction (<30%) has been shown to be a predictor of increased operative mortality (10, 11). In the AE, recognized preoperative risk factors are assigned numerical values from 0 to 3. These are added together to give a total score that predicts the risk of operative mortality as a percent. At the current time, the AE is occasionally used in Kenya; more commonly, the presence of an Ejection fraction of less than 30% is used to inform patients that they are at an increased risk for operative mortality. This study aimed to assess the applicability of the AE to a population of cardiac surgical patients in Nairobi, Kenya.

Methods

A retrospective study was carried out at the Kenyatta National Hospital (KNH) Nairobi (Kenya) between 1st January 2011 and 31st December 2015. The operative register in the cardiothoracic surgery theatre was reviewed. All patients aged 18 years or older who had valve surgery, coronary artery surgery or both on cardio-pulmonary bypass at KNH during the study period were included; a consecutive sample of patients was taken. Patient's case notes with incomplete data were excluded. Risk factor prevalence was compared with that of the population used to derive the Euroscore. The 'Chi-square' test was used for categorical variables and the Student's t-test for means. The AE was calculated and its discrimination was determined by receiver operator curve (ROC) analysis. The study was carried out following approval by the Kenyatta National Hospital/University of Nairobi- Ethics and Research Committee.

Results

Two hundred and two patients were identified from the cardiothoracic surgery theatre registry; 148 (73%) case files were available, 109 (54%) of which contained adequate information and were included in the study. The mean age was 35.1(+/- 11); the in-hospital (operative) mortality was 13.8%. Table 1 summarizes the risk factor prevalence in the studied patients. Following risk factor analysis (see table 2), only double valve replacement (DVR) was found to be a risk factor for operative mortality; odds ratio 5.98 (1.83 to 19.49; p-value 0.003). The area under curve (AUC) for the AE score was 0.59; the curve is shown in figure 1; the grey lines represent the 95% confidence interval. Table 3 compares the risk factor profile of our study population with that of the Euroscore population and significant issues are discussed in the discussion section.

Table 1: Risk Factor Prevalence

Risk factor	Prevalence
Pulmonary hypertension (PHT)	58.7
Left ventricular dysfunction (LVD)	37.6
Double valve replacement (DVR)	17.4
Impaired renal function (IRF)	1.8
Female Isolated coronary artery	67.9
Coronary artery bypass grafting (CABG)	4.6

Table 2: Outcomes in DVR and Non-DVR patients

	Alive	Dead
DVR	12	7
Non-DVR	82	8

DVR, Double Valve Replacement

Table 3: Comparison of Risk Factor Prevalence.

Risk factor	Prevalence in this study (%)	Prevalence in Euroscore population (%)	p-value
N	109	19030	
Age*	35.1(+/- 11)	62.5(+/- 10.7)	<0.001
PHT	58.7	2.0	<0.01
LVD	37.6	31.4	0.11
DVR	17.4	14.0	0.30
IRF	1.8	1.8	0.98
Female	67.9	27.8	<0.001
Isolated CABG	4.6	65.0	<0.001

CABG, Coronary Artery Bypass Grafting; DVR, Double valve Replacement; IRF, Impaired Renal Function; LVD, Left Ventricular Dysfunction; N, Number of patients studied; PHT, Pulmonary Hypertension   *mean in years
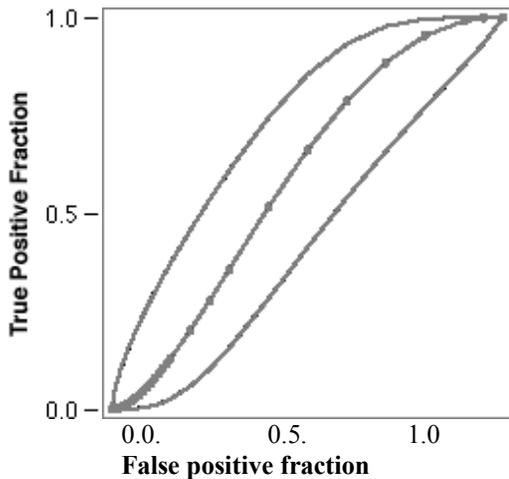
Figure 1: Receiver operator curve for AE; AUC is 0.59

Discussion

Useful risk scores must have good discrimination and calibration (12). Discrimination refers to a scores ability to distinguish between patients at low and high risk for a particular outcome. For example, if the majority of deaths occurred in patients that the model identified as high risk, the model is said to have good discrimination. On the other hand, if most deaths occur in patients that the model identified as low risk, the model has poor discrimination (13). Calibration refers to a score's ability to predict risk in a group of patients. For example, if the model predicts that mortality in 1,000 patients would be 10% and if the observed mortality is 10% (or close to this), the model is well calibrated (13). There is a limit to the ability of regression analysis to identify risk factors (14). Increasing the number of variables may improve score performance, however, including too many variables may lead to "over fitting" and thereby reduces the number of populations in which the score may be applicable (15). Table 3 compares the prevalence of risk factors in our study patients and in those used to develop the AE. There are three striking features about this comparison: our patients were significantly younger, more of our patients had pulmonary hypertension (PHT) and very few of our patients had isolated coronary artery bypass graft (CABG). As far as patient age is concerned, the AE assigns our patients and those used to develop the AE, the same score. It means that for purposes of the AE, patient age was not different between the two groups. The highest value allocated to a particular risk factor in the AE is 3 and the lowest is 1. Poor left ventricular function is

allocated 3 and PHT is allocated 2. Our patients had a significantly higher prevalence of PHT and so had a significantly higher prevalence of a 'moderate' risk factor for operative mortality. The majority of our patients had a procedure other than isolated CABG and were allocated a score of 2 for this parameter. In contrast, the majority of patients used to create the AE had isolated CABG and this was not considered a risk factor for mortality and was allocated a score of "0". This means that the majority of our patients had this additional moderate risk factor for mortality. Taken together the majority of our patients had 2 additional moderate risk factors for operative mortality compared to the patients used to develop the AE. This may explain why the AUC was low (0.59). It may well be that these 2 additional risk factors in the majority of our patients, rendered our population too different for the AE to be valid in our population. As the discrimination of the AE in our patients was poor (AUC <65) it was deemed pointless to assess the calibration so this was not performed.

Surprisingly, the prevalence of poor left ventricular function was not significantly different in the 2 populations. This was surprising because a common hypothesis in our unit for the higher operative mortalities is that our patients presented later and so are more likely to come to surgery with more advanced disease manifesting as a greater prevalence of poor left ventricular function. Our study population is quite small and almost half of the patients operated on during the study period could not be included in the study. Despite this our findings do raise the possibility that the AE may not be applicable to our patients. In addition, it was apparent from this study that the quality of our medical record keeping is poor. This is a significant problem as complete records are necessary to develop our own risk stratification system.

In our study, DVR was a risk factor associated with increased operative risk. DVR requires a longer time on cardiopulmonary bypass (CPB) and extended CPB time is a traditional risk factor for operative mortality. However, the prevalence of DVR in our patients was not significantly different from that of those from whom the AE was developed.

Conclusion

Our study demonstrates that our patients may have had significantly more risk factors for operative morality than the patients used to develop the AE. This may explain the poor discrimination of the AE in our patients. It would seem reasonable to develop our own local risk stratification system. As DVR was found to be a risk factor for mortality in our patients, our local risk score should incorporate this risk factor. It would be prudent to allocate resources towards developing a cardiac surgery data base to ensure that adequate data is available for risk stratification score development.

References

1. Greenhalgh T, Howick J, Maskrey N. Evidence Based Medicine: A Movement in Crisis? BMJ 2014; 348:g3725.
2. Kelly MP, Capewell S. Relative Contributions of Changes in Risk Factors and Treatment to the Reduction in Coronary Heart Disease Mortality. Health Development Agency, 2004.
3. Lau BD, Haut ER. Practices to Prevent Venous Thromboembolism: A Brief Review. BMJ 44 Qual Safety 2014; 23:187-95.
4. Bridgewater B, Grayson AD, Jackson M, et al. Surgeon Specific Mortality in Adult Cardiac Surgery: Comparison between Crude and Risk Stratified Data. BMJ 2003; 327:13–7.
5. Yap CH, Reid C, Yii M, et al. Validation of the Euroscore model in Australia. Eur J Cardiothoac Surg. 2006; 29(4):441- 6.
6. Parsonnet V, Dean D, Bernstein AD. A Method of Uniform Stratification of Risk for Evaluating the Results of Surgery in Acquired Adult Heart Disease. Circulation 1989; 79:I3-12.
7. Bridgewater B, Neve H, Moat N et al. Predicting Operative Risk for Coronary Artery Surgery in the United Kingdom: A Comparison of Various Risk Prediction Algorithms. Heart 1998; 79: 350-5.
8. Tiveron MG, Bomfim HA, Simplicio MS et al. Performance of Inscor and Three International Score in Cardiac Surgery at Santa Casa de Marila. Braz J CardiovascSurg. 2015; 30(1):1-8.
9. Grant SW, Hickey GL, Dimarakis I et al. How Does the Euroscore II Perform in UK Cardiac surgery. An Analysis of 23740 Patients from the Society of Cardiothoracic surgery in Great Britain and Ireland National database. 2012; 98(21): 1568-72.
10. Roques F, Nashef SA, Michael P, et al. Risk Factors and Outcome in European cardiac surgery: Analysis of the Euroscore Multinational Database of 19030 Patients. Eur J Cardiothorac Surg. 1999; 15(6):816-22.
11. Prins C, I de Villiers J, Botes L, et al. Cardiac Surgery Risk Stratification Models. Cardiovasc J Afr. 2012; 23: 160-4.
12. Nashef SA, Roques F, Hammill BG, et al. Validation of European System for Cardiac Operative Risk Evaluation (Euroscore) in North American cardiac surgery. Eur J Cardiothorac Surg. 2002; 22:101-5.
13. Mejía OAV, Lisboa LAF, Puig LB, et al. InsCor: A Simple and Accurate Method for Risk Assessment in Heart Surgery. Arq Bras Cardiol. 2013; 100(3):246-54.
14. Diamond GA. Future Imperfect: The Limitations of Clinical Prediction Models and the Limits of Clinical Prediction. J Am Coll Cardiol 1989; 14: 12A-22A.
15. Ranucci M, Castelvecchio S, Menicanti L, et al. Accuracy, Calibration and Clinical Performance of the Euroscore: Can we Reduce the Number of Variables? Eur J Cardiothorac Surg. 2010; 37(3):724-9.